Aalto University
School of Science
Master's Programme in Industrial Engineering and Management

Henri Huttunen

The role of data in firm performance:

A techno-economic view

Master's Thesis
Espoo, July 25, 2019

Supervisor and advisor: Timo Seppälä, Professor of Practice

Aalto University
School of Science
Master's Programme in Industrial Engineering and Management

ABSTRACT OF MASTER'S THESIS

Author:	Henri Huttunen		
Title: The role of data in firm performance: A techno-economic view			
Date:	July 25, 2019	**Pages:**	IX + 84
Major:	Strategy and Venturing	**Code:**	SCI3050
Supervisor and advisor:	Timo Seppälä, Professor of Practice		

Data is an increasingly discussed topic in both academia and business; industries considered as traditional have had to embrace the digitalization and the subsequent transformation in the past decades. The question remains, what kind of improvements could companies realize if they were to utilize data to its full extent in their business.

The evolution of information technologies and the progression of digitalization has created the opportunity to expand product and service modularity from hardware to software and platforms; nowadays, many successful technology companies create products that are further improved by external developers creating new applications and features to the products, and they can be made readily available to anyone with little cost.

This study links the literature from information technologies to the economic benefits enabled by data to form a holistic overview on how to review the impact of data on firm performance. The theoretical framework is then validated by triangulation of quantitative data of Finnish companies and expert interviews.

The benefits of data can be divided into operational efficiencies and strategic opportunities—cost reductions and new products and services—and into internal and external sub-dimensions. In order to achieve any external benefits, data must be shared with external parties and the development of data sharing technologies from EDIs to APIs has significantly lowered the barriers to realizing these efficiencies and opportunities.

Ultimately, the results of this study give business managers the lenses to assess what kind of benefits their company has already realized, and what kind of opportunities would be available if their data strategy was modified to harness the opportunities enabled by the latest technologies.

Keywords: data sharing, api, big data, analytics, modularity
Language: English

Aalto-yliopisto
Perustieteiden korkeakoulu
Tuotantotalouden maisteriohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Henri Huttunen		
Työn nimi: Datan rooli yritysten suorituskyvyssä: Tekno-ekonominen näkökulma			
Päiväys:	25. heinäkuuta 2019	**Sivumäärä:**	IX + 84
Pääaine:	Strategy and Venturing	**Koodi:**	SCI3050
Valvoja ja ohjaaja:	Professori Timo Seppälä		

Datasta keskustellaan entistä enemmän sekä akateemisissa että liiketoiminnallisissa yhteyksissä; jopa perinteisinä nähdyt toimialat ovat joutuneet omaksuneen digitalisaation ja sen myötä tulleet muutokset liiketoimintaan viimeisten vuosikymmenien aikana. Millaisia tehokkuuksia yritykset voisivat saavuttaa mikäli ne hyödyntäisivät dataa parhaimmalla mahdollisella tavalla liiketoiminnassaan?

Informaatioteknologioiden sekä digitalisaation kehitys ovat luoneet mahdollisuuden laajentaa tuotteiden ja palveluiden modulaarisuuden käsitettä fyysisistä tuotteista ohjelmistoihin ja alustoihin; nykypäivänä monet menestyneet teknologiayritykset kehittävät tuotteita jotka kehittyvät entisestään kun ulkopuoliset kehittäjät luovat niille uusia tuotteita ja ominaisuuksia jotka voidaan jakaa käyttäjille mitättömin kustannuksin.

Tämä tutkimus yhdistää kirjallisuuden informaatioteknologioiden tarjoamista mahdollisuuksista datan hyödyntämisen taloudellisiin seuraamuksiin muodostaen kattavan mallin siitä, kuinka datan vaikutusta yritysten suorituskykyyn voidaan tarkastella. Kehitetty teoreettinen viitekehys validoidaan yhdistämällä tilastollista dataa Suomalaisista yrityksistä asiantuntijoiden haastatteluihin.

Datan mahdollistamat hyödyt voidaan jakaa operatiivisiin tehokkuuksiin sekä strategisiin mahdollisuuksiin—kulujen vähennyksiin sekä uusiin tuote- ja palvelumahdollisuuksiin—ja sisäisiin sekä ulkoisiin ulottuvuuksiin. Jotta ulkoisia hyötyjä voitaisiin saavuttaa, dataa täytyy jakaa ulkopuolisten toimijoiden kanssa. Datan jakamisen teknologioiden kehitys EDI-järjestelmistä API-ratkaisuihin on merkittävästi madaltanut kynnystä näiden tehokkuuksien sekä mahdollisuuksien saavuttamiselle.

Tämä tutkimus antaa yritysten johtohenkilöille tavan arvioida millaisia datan hyötyjä yritys on jo saavuttanut sekä millaisia hyötyjä olisi vielä saavutettavissa mikäli yrityksen datastrategiaa muokattaisiin hyödyntämään uusimpien teknologian tarjoamia mahdollisuuksia.

Asiasanat:	datan jakaminen, rajapinnat, massadata, analytiikka, modulaarisuus
Kieli:	Englanti

Acknowledgements

I never knew what to expect from writing a thesis but certainly I did not expect it to be an experience during which I would learn so much. However, I am very pleased of how everything turned out as I would rather constantly learn new things than keep exploiting only the knowledge and know-how that I have gathered previously.

First, I would like to express my gratitude to my advisor and supervisor Timo Seppälä with whom I have had the pleasure to work with since completing my Bachelor's thesis under his supervision. He introduced me to the world of platform economy and the opportunities digitalization offers in the business-to-business context, and he has been extremely helpful and resourceful in supervising this thesis.

Additionally, The Research Institute of Finland and Aalto University deserve my deepest gratitude for contracting this thesis as a part of an industrial data project funded by The Finnish Innovation Fund Sitra. It has been both interesting and a thorough learning experience to see the academic work from the perspective of a research assistant.

Finally, I am grateful for the people who have been around me during the process; my girlfriend Rosa and my friends—particularly Kasper—have had to bear with me through the various stages of frustration and joy.

Espoo, July 25, 2019

Henri Huttunen

Abbreviations and Acronyms

API	Application Programming Interface
BI&A	Business Intelligence and Analytics
CAGR	Compounded Annual Growth Rate
VAN	Value Added Network
EDI	Electronic Data Interchange
I-EDI	Internet Electronic Data Interchange
IOS	Interorganizational Information System
IoT	Internet of Things
M2M	Machine-to-Machine
PSD2	Payment Services Directive 2015/2366
XML	eXtensive Markup Language

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Data is an increasingly prevalent topic in business — in 2012, 12% of Fortune 1000 companies had a Chief Data Officer, and in 2018 67.9% of surveyed companies (n=65) reported having a CDO[1]. Although the small sample of the Fortune 1000 companies raises the possibility of a sample bias, the trend is still noteworthy. While the most advanced manufacturers have started gathering data of their processes by integrating connectivity into their products and machinery, the vast majority of the data never leaves their internal systems — such system should therefore rather be called an "Intranet of Things" than "Internet of Things" (Fitzgerald, 2013). Such data could potentially be very valuable, and additional benefits could be achieved if it was monetized or shared accordingly.

Historically, many of the benefits related to data—and the underlying mechanisms—have been present in other forms. Recent technologies have enabled product modularity to move from the hardware, such as machines and computers (Henderson and Clark, 1990; Schilling, 2000), to the services and content of the products as outlined by Yoo, Henfridsson, and Lyytinen (2010) and Ghazawneh and Henfridsson (2010). Voss and Hsuan (2009) expanded modularity from product systems to service systems and argued that efficiencies of modularity can also be realized in service processes, and recently the discussion has shifted to platform modularity (Gawer and Cusumano, 2008; Gawer, 2014). To illustrate the change, personal computers are physically modular (various components e.g. graphics cards can be manufactured by anyone and attached to the computer via standard connections) and modern smartphones' configurations are modular via software (every person can have a smartphone that it specifically tailored to the user's needs by installing the specific applications and software modifications desired). Software modular-

[1] `https://www.forbes.com/sites/insights-intelai/2019/05/22/rethinking-the-role-of-chief-data-officer/` (Accessed 18.06.2019)

ity discussion is now closely tied into platform-centric discussion, because the physical products can create platforms and ecosystems around them attracting external parties to develop these software modules (Katz and Shapiro, 1994; Gawer and Cusumano, 2008; Tiwana, Konsynski, and Bush, 2010; Tiwana and Konsynski, 2010).

To facilitate the cooperation and coopetition, boundary resources[2] have to be set that enable the third parties can use in their development (Gawer, 2009; Ghazawneh and Henfridsson, 2010). When addressing the role of data in firm performance, the companies' data sharing technologies and mechanisms act as boundary resources, as they serve as the interface between the company and the application developers (Ghazawneh and Henfridsson, 2010).

From a technology standpoint, ample literature exists on various information technologies and their impact on firm performance; the benefits enabled by adopting EDI technologies that allowed instantaneous information and data sharing within companies and with external parties (Pfeiffer, 1992; Arunachalam, 1995; Mukhopadhyay, Kekre, and Kalathur, 1995), how the integration and adoption of web based systems enabled even further benefits (Stefansson, 2002; Zhu et al., 2006; Huang, Janz, and Frolick, 2008), and how the emergence of APIs—that allow thousands of networked systems to communicate with each other in real time—and the related "API economy" (Moilanen, Niinioja, Seppänen, and Honkanen, 2019) have disrupted businesses and created new opportunities altogether (Jacobson, Brail, and Woods, 2011; Smith, Ofe, and Sandberg, 2016).

Having access to your company's operational data, and sharing it within their supply chain has led to superior operational efficiency as has been noted in the literature from 1990s (Stevens, 1989; Tan, Kannan, and Handfield, 1998; Gavirneni, Kapuscinski, and Tayur, 1999; Cachon and Fisher, 2000; Lee, So, and Tang, 2000; Lau, Mak, and Huang, 2003). More recently, the strategic value of data has also been recognized, particularly as big data is becoming accessible to virtually any party; Constantiou and Kallinikos (2015) and Woerner and Wixom (2015) discuss the strategic relevance of data particularly in decision-making, and additional opportunities related to complementary innovations (Zhu and Liu, 2018) and data monetization (Liozu and Ulaga, 2018) are actively being discussed.

The purpose of this study is to form a theoretical framework for assessing the business opportunities of properly utilizing data, analytics, and data sharing practices in companies. In practice, this study takes and abductive approach by combining literature from information technology and business,

[2]For more discussion regarding boundary resources, see section 2.1.

and triangulation of expert interviews and statistics[3] of Finnish companies to create and validate a techno-economic view of the matter.

The research questions have been formulated as follows:

1. *What kind of performance-related benefits does data enable in companies with modern technologies?*
    (a) *How has the evolution of information technology affected the sharing and use of data and information?*
    (b) *What kind of benefits does the use and sharing of data entail in companies?*
    (c) *How could the benefits of data-related activities be quantified?*

Through these research questions, the aim is to formulate a hypothesis for quantifying the impact of data on firm performance — considering the firm's level of technology adoption and the data exploitation practices in place. Quantifying the balance sheet value of data is left outside of the scope of this study[4]; the focus remains on assessing how data impacts the firm's financial performance i.e. the revenue, gross margins, or market capitalization.

In order to answer these questions, the technological readiness and adoption were studied with historical data of Finnish companies by Statistics Finland, with some additional data from The Research Institute of the Finnish Economy. Several high-profile experts with diverse set of business and technological background were then interviewed in order to complement the statistics and understand how the exploitation and the impact of data has developed in their experience.

The results of the study show that the evolution of technology has made the benefits available to an increasing portion of all companies, and data has indeed become an important aspect of businesses across industries. Companies are integrating analytics and data-driven decision making into their companies, but the motivation remains largely in achievable operational efficiencies with some actions taken to improve revenues by understanding the customers better. Therefore, strategic opportunities remain largely untapped, although interviews suggest that companies are actively exploring their options — without hard evidence of potential benefits and without proper capabilities, very few companies are actively pursuing external strategic opportunities enabled by opening APIs and allowing external access to data.

[3]Statistics provided by Statistics Finland, for more detail, refer to chapter 3.

[4]Valuing information technology related intangible assets is an extremely complex issue, see Brynjolfsson, Hitt, and Yang (2002) for more discussion on the topic.

Structure of the thesis

In chapter 2, the previous literature will be assessed in two parts; the technological evolution of data sharing technologies to understand what role the development of information technology has had on firms' performance in this context, and the identified economical benefits of data economy.

The two viewpoints of literature will be used to create a hypothesis for a framework tying them together. Chapter 3 outlines the methodology used to refine and validate the hypothesis — this chapter also includes detailed description of the empirical data used in this study.

Empirical data will be analyzed and discussed in chapter 4 while the discussion of the results, their implications, and the limitations of the study are assessed in chapter 5. In chapter 6, the most important insights are summarized in conclusion.

Chapter 2

Literature review

This literature review consists of two main parts that are preluded by a shorter section outlining the theoretical context of the study. The first main section will provide the technological viewpoint—assessing how did we get here—to this study as the evolution of data sharing mechanisms will be assessed from the first electronic data interchange systems to APIs, that have enabled real time data sharing between a virtually limitless amount of parties. In the second section, the benefits of data will be reviewed to establish business perspective of the impact of data to companies' businesses by considering the various ways data can be used and what benefits can be associated with it.

2.1 Theoretical context of the study

Although the study is focused on the role of data on firm performance, it is important to outline the wider context in which the study takes place. Assessing the evolution of the data sharing mechanisms in section 2.2 is relevant due to two distinct perspectives; 1) the more sophisticated data sharing mechanisms have enabled product modularity on the software level, when relevant data and software can be made available to external parties, and 2) the interfaces of these technologies function as boundary resources for external parties to use. In this section, the key concepts of these two perspectives are addressed briefly.

Modular product design

Modular product design stems from an article by Starr (1965) in which he proposes a new definition of modular products that are designed as a set of independent modules, which can be reused and interchanged to maximize product variety. Literature suggests that modular product design has positive influence on the products' performance; Antonio, Yam, and Tang (2007) show that product modularity influences the capabilities of delivery, flexibility, and customer service of which the first two positively relate to product performance. Originally product modularity was used in hardware components of various machines and computers (Henderson and Clark, 1990; Schilling, 2000), but with the advancement of information technologies, modular product design has—following a layered architecture model by Yoo et al. (2010)—moved from the devices to services and content as well. While Sanchez and Mahoney (1996) use aircraft, automobiles, and consumer electronics as prime examples of product modularity, Gawer and Cusumano (2008) and Tiwana and Konsynski (2010) refer to product modularity in the context of platforms and software modularity, supporting the evolution from device layer to content layer.

Modular design is hence interconnected with platform literature — Gawer (2014) outlines three types of platforms; 1) internal platforms, which have closed interfaces for use within the company, 2) supply-chain platforms in which interfaces are selectively open for partners in the company's supply chain, and 3) industry platforms, which have open interfaces available to all external parties. Modular design in the platform's technological architecture is key in order to exploit the capabilities accessible through the external parties (Gawer, 2014).

Furthermore, software development is increasingly shifting toward platform-centric ecosystems (Katz and Shapiro, 1994) in which the platform provides an interface[1], often an API, to which external modules connect to and add functionality (Tiwana et al., 2010). In this study, the various assessed data sharing technologies are interfaces in the same context.

Before platform modularity became the focus of modularity discussion, some researchers, such as Voss and Hsuan (2009) and Tuunanen and Cassab (2011), examined service modularity. Services are obviously very different from physical products, and interfaces in modular service design can be e.g. people and information (Voss and Hsuan, 2009) and it can generate market impact efficiently through innovative offerings by reusing and varying existing services (Tuunanen and Cassab, 2011).

As for the benefits of modern modularity, Tiwana (2008) proposes that modularity decreases the need for control and actually increases performance

[1]Interfaces are linkages shared among components and can be considered as "an elaboration of the physical architecture that comprises a minimal set of rules governing the arrangement, interconnections, and interdependence of the elements" (Voss and Hsuan, 2009).

in alliances, and uses Amazon and Google as examples[2]. IT modularity also increases IT agility as the systems are more interoperable and easier to integrate (Tiwana and Konsynski, 2010).

Tiwana and Konsynski (2010) define IT architecture modularity as "the degree of decomposition of an organization's IT portfolio into loosely coupled subsystems that communicate through standardized interfaces" that consists of two dimensions; IT architecture loose coupling and IT standardization. By loose coupling, they refer to the degree to which the components or applications in an organization's IT architecture are designed such that internal changes in one application do not affect the behavior of others, and standardization refers to organizationwide standards and policies that define how the IT systems connect and interoperate with each other (Tiwana and Konsynski, 2010).

Modularity has progressed from physical products to services, and furthermore into software platforms. This has made the realization of its benefits more easily accessible to a wider range of companies, and it is largely driven by the evolution of information technologies and discussed in section 2.2.

Data sharing technologies as boundary resources

In data sharing context, technologies enabling the transmission and sharing of data can be considered technological boundary resources (as discussed by Ghazawneh and Henfridsson (2010)) through which companies can collaborate with each other. In figure 1 the hierarchy of the terminology is laid out with selected references for each level.

[2]Google provides a software platform, Android, which is highly modular and developers can freely develop applications and modifications to the software. They have very little control over the developed software, but their software is very successful.

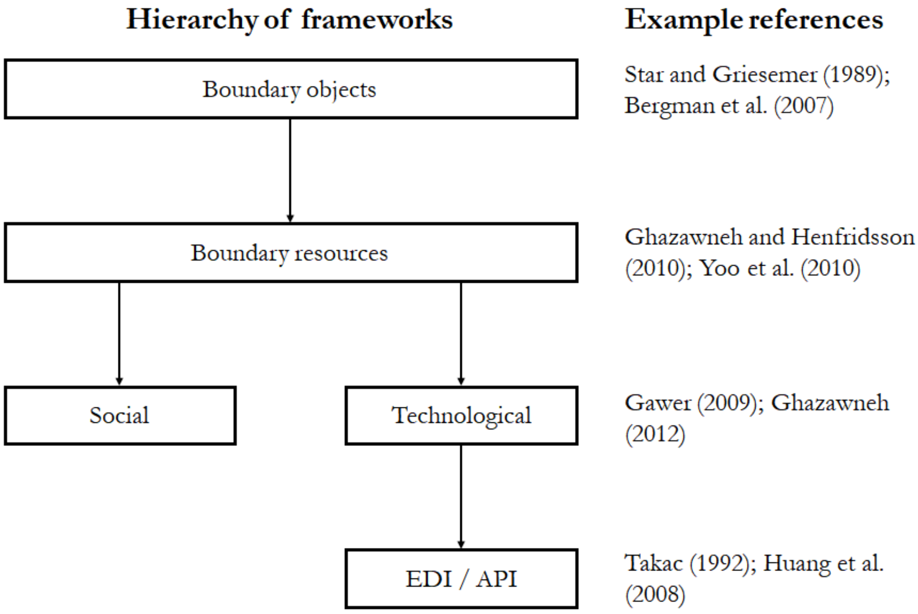

Figure 1: Hierarchy of terminology

The foundation of boundary resources stems from sociological studies where Star and Griesemer (1989) introduce "boundary object" as a theoretical tool for conceptualizing how various actors with conflicting objectives can cooperate in a project. Examples of boundary objects are shared repositories, ideal types, coincident boundaries, and standardized forms (Star and Griesemer, 1989). In the context of software platforms, Ghazawneh and Henfridsson (2010) define "boundary resources" on the basis of boundary objects — they are software tools and regulations that serve as the interface between the platform owner and the application developers. Furthermore, boundary resources can be broken into social (e.g. incentives, intellectual property rights, guidelines) and technological (e.g. software development kits, application programming interfaces) (Gawer, 2009; Ghazawneh, 2012). In this study, the assessed technologies are considered as such technological boundary resources that help facilitate the data exchange between companies.

2.2 Evolution of data sharing mechanisms

Information and data[3] have been used, shared, and traded between individuals and organizations throughout the history, but it has undergone drastic

[3] In order to exchange information electronically, it must be digitized — hence, data can simply be defined as digitized information.

changes in the past decades due to digitalization. Takac (1992) points out that data was already moved electronically in the late 1800s when the US railroad was using the telegraph to identify shipments and their contents. The development of data exchange has accelerated drastically due to the growth of digitalization; in the year 2000, approximately a quarter of all information was stored digitally, and the same number was 98% in 2013 (Mayer-Schönberger and Cukier, 2013). The evolution of data sharing technologies has fueled this growth of data due to the increasing amount of devices and parties connected to each other, and a report[4] estimates that by 2025 every person on Earth will have access to the internet and can both share their own data, and access the openly available data and information in the world.

To be able to assess the potential and impact of this paradigm, it is necessary to outline the evolution of data sharing mechanisms and technologies (in the business to business context) to understand how each step of the evolution of data sharing has impacted the companies' performance.

To further underline the importance of the evolution of data sharing technologies, World Economic Forum[5] published an article in 2019 in which similar technologies are shown to improve companies cumulative capabilities as they have evolved. However, it is important to note that these technologies are adopted and used for more use-cases than exchanging information and data with external parties—often for internal purposes—and therefore the impact of these technologies must be assessed holistically. Additionally, several academic studies suggest that using information technologies creates value for companies but due to the complexity of the issue, it is difficult to quantify the exact impact (Brynjolfsson and Hitt, 2000). Hence, it is important to assess the drivers through which these technologies create the value to the companies, rather than focus on the technologies alone, and link it with the benefits associated with analyzing the data or exchanging data and information outside the company borders.

2.2.1 Electronic data interchange

Electronic data interchange (EDI), a type of interorganizational information system (IOS) (Takac, 1992), can be defined as "process of computer to computer, business to business data transfer of repetitive business processes involving direct routing of information from one computer to another without human interference, according to predefined information formats and rules"

[4] `https://www.mirror.co.uk/news/world-news/everyone-earth-internet-access-2025-3812415` Accessed 18.06.2019

[5] `https://www.weforum.org/agenda/2019/05/we-need-to-measure-innovation-better-heres-how-to-do-it/` Accessed 18.06.2019

(Narayanan, Marucheck, and Handfield, 2009), or more generally as systems that firms or organizations use to exchange information and data within or between each other electronically. They were essentially the first information technology systems that connected companies to each other and allowed the exchange of information and data before the internet. Arunachalam (1995) lists ordering and paying for goods from suppliers, arranging transportation with carriers, receiving orders from customers, invoicing customers and collecting payments from customers as general use cases of EDI — all cases in which information was previously transferred manually via e.g. a telephone. A more formal definition by Pfeiffer (1992) requires EDI to possess four features:

1. It must have a least two organizations in a business relationship as users
2. Data processing tasks pertaining to a transaction at both (all) organizations must be supported by independent applications systems
3. The integrity of the data exchange between application systems of trading partners must be guaranteed by agreements concerning data coding and formatting rules
4. Data exchange between the application systems must be accomplished via telecommunication links

Although EDIs have been widely researched as a means to improve information flow and communications between a company and its partners, especially in supply chain management (e.g. Mukhopadhyay et al. 1995), the literature remains inconclusive regarding the exhaustive list of benefits gained from its usage (Ahmad and Schroeder, 2001; Barratt and Oke, 2007). In their literature review, Narayanan et al. (2009) found discrepancies in the literature between the reported benefits and their impact; some companies were able to realize significant savings while others failed. The hypothesised reasons for such discrepancies varied from differences in industries, business models, and competencies.

As a delivery platform, EDI typically used a privately owned value added network (VAN), which hosted a VAN mailbox each EDI adopter subscribed to in order to exchange messages with other VAN subscribers as illustrated in figure 2 (Arunachalam, 1995; Bury, 2005; Zhu, Kraemer, Gurbaxani, and Xu, 2006) — the proprietary nature of the VAN required all parties to develop their own compatible systems, making the system very costly to set up, maintain, and use (Bednarz, 2004; Huang et al., 2008). In practice, use of a VAN creates a single electronic pipeline to a client: suppliers connect to the VAN, send their information to it, and the VAN delivers it to buyers (Bury, 2005). Hence, the networks remained relatively small with only the large companies and their key partners communicating with each other via the VAN.

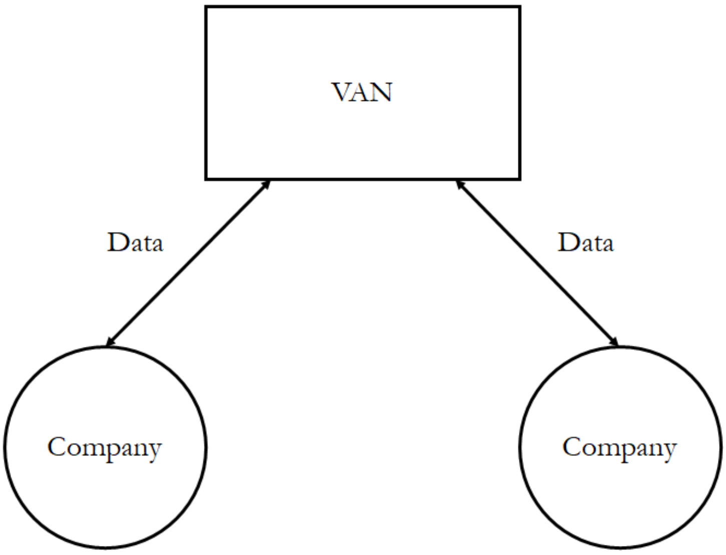

Figure 2: Illustration of a VAN based EDI connection

Like many other technologies, EDIs were initially used primarily by large companies due to high costs associated with adopting the technology and the lack of required technological skills posing too high barriers for smaller companies (Pfeiffer, 1992; Arunachalam, 1995; Iacovou, Benbasat, and Dexter, 2006). Other perceived barriers to adoption were simply lack of awareness of its benefits, lack of partners' participation, lack of standard formats, and incompatibility of existing hardware or software (Arunachalam, 1995).

EDIs started appearing in academia from the late 1980s (e.g. Robinson and Stanton 1987) and started gaining traction in the 1990s when Emmelhainz (1992) wrote an EDI management book, Takac (1992) addressed the potential performance improvements, and later Reekers (1994) conducted a survey on EDI use and adoption that has since been widely cited in later studies. Of his respondents, the earliest companies had implemented EDIs in 1978 and Zhu et al. (2006) found that the adoption started around 1970. With the increasing penetration of information technology, EDI adoption grew rapidly, especially in companies in the USA and Europe. Their adoption and impact also started to become an important topic in academic research after 1990 (see e.g. Reekers 1994; Mukhopadhyay et al. 1995; Arunachalam 1995, 1997).

While an exhaustive and exact analysis on the benefits of EDI is too complex to conduct, table 2 summarizes the main realized benefits of EDI found in literature. The results align well with the purpose of EDI — its use allowed companies to streamline their manual processes with their suppliers and partners, and thus reduce costs associated with the labor (Reekers, 1994; Mukhopadhyay et al., 1995; Arunachalam, 1997; Iskandar et al., 2001). The other benefits are closely related to this main result; gross margins can be increased due to the reduced costs associated with information inefficiencies, and communications are improved due to the streamlined processes and integrated information databases. As a practical example, General Motors moved from paper checks to paperless EDI payments and the projected annual savings were approximately $1.3 billion from that project alone (Arunachalam, 1995). Other benefits such as reduced inventory costs (Mukhopadhyay et al., 1995; Vijayasarathy and Tyler, 1997) or improved customer service (Reekers, 1994; Arunachalam, 1997; Iskandar et al., 2001) were also identified, but such findings can be categorized under the three main benefits. However, the results of Reekers (1994) indicate that companies primary reason for EDI adoption was not the reduction in costs, but rather the improved communications and cost reductions were merely a byproduct. Similarly, the results from Arunachalam (1995) indicate that the main reasons for adoption were actually customer requests, remaining competitive, and better customer service — two of the top reasons are arguably related to the improved communications referred to by Reekers (1994).

Perspective on EDI	Description	Selected references
EDI adoption and management	EDI initially had very high adoption costs and was primarily accessible by large companies. Additionally, a lack of technical skills within organizations was hindering both adoption and use.	Reekers (1994); Crum, Premkumar, and Ramamurthy (1996); Arunachalam (1997); Vijayasarathy and Tyler (1997); Lee et al. (2000); Iskandar, Kurokawa, and LeBlanc (2001); Iacovou et al. (2006)
EDI's impact on performance	Literature is inconclusive in quantifying the exact impact of EDI adoption on firm performance. However, several studies suggest benefits driven by improved processes in information handling and supply chain management	Takac (1992); Reekers (1994); Mukhopadhyay et al. (1995); Arunachalam (1995); Prosser and Nickl (1997); Lee et al. (2000); Ahmad and Schroeder (2001); Iacovou et al. (2006)
EDI as a technology	EDI systems were operated on value added networks (VAN) and such systems were proprietary and had to be set up individually with all parties in the network.	Arunachalam (1995); Vijayasarathy and Tyler (1997); Lee and Lim (2003); Zhu et al. (2006); Huang et al. (2008)

Table 1: Summary of the reviewed EDI literature

Benefit	Selected drivers	Selected references
Competitive benefits	Increased gross margin, improved product/service quality	Reekers (1994); Crum et al. (1996); Vijayasarathy and Tyler (1997)
Cost-saving benefits	Lower general management costs, decreased inventory costs, staff reductions, streamlined processes	Reekers (1994); Mukhopadhyay et al. (1995); Arunachalam (1995, 1997); Iskandar et al. (2001)
Supply chain communication and coordination benefits	Improved communications with suppliers, integrated information database	Reekers (1994); Banerjee and Golhar (1994); Arunachalam (1997); Iskandar et al. (2001)

Table 2: New benefits enabled by EDI

The identified benefits of EDI all relate to improving the companies' current processes, and largely target specific functions either within the company, or between specific external stakeholders. It is important to note, that the EDI systems between organizations were quite rigid; as Narayanan et al. (2009) state, the information exchange is done according to predefined rules and formats and was typically an electric dyad integrating a single buyer and supplier. However, as the EDIs started transitioning into internet their adoption and scope grew (Bednarz, 2004; Narayanan et al., 2009) which is discussed in the following section.

2.2.2 Open-standard interorganizational information systems

Internet-based IOS technologies started appearing with the widespread adoption of the internet in the late 1990–era and the internet provided a new technology for doing EDI (Stefansson, 2002). They were introduced to upgrade or substitute traditional IOS systems, such as VAN-based EDIs, and in 2000 46% of the 50 Fortune 1000 firms surveyed planned to use Internet EDI (I-EDI) in one form or another (Stefansson, 2002; Huang et al., 2008). Fuelling the adoption, the I-EDIs were capable of tackling most of the barriers to EDI adoption companies perceived around 1995 (Arunachalam, 1995). In the literature I-EDIs are also discussed as open-standard interorganizational information systems and Zhu et al. (2006) define them as the kind of interorganizational systems that use open standards (e.g. TCP/IP[6] as the communication protocol, and XML[7] as the data standard) and are built upon the open internet for information exchange and business-to-business transactions. It should be noted, that I-EDI was not a system in itself; it is a term to describe systems operating under the aforementioned conditions. For example, extranets were used by a company and its business partners by allowing external parties access to the company's internal information systems (Watson, 1999; Stefansson, 2002).

Table 3 outlines some of the the key differences between the VAN-based EDIs and the new, internet-based IOS, but arguably the most important change was the reduced cost of adoption from approximately \$100 000 to \$10 000 (Werner, 1999) and the widely expanded accessibility over the web. This allowed much smaller companies to adopt EDIs and the benefits became widely accessible further down the supply chains (Huang et al., 2008). Combined with the fact that the internet and third-party transaction integration services provided companies with increased possibilities to network with supply chain partners (Auramo, Kauremaa, and Tanskanen, 2005), the information sharing technologies became much more common in the era of internet and I-EDIs.

Network effects

Network effects recognized by Zhu et al. (2006) refer to dynamics in which the parties on a platform, in this case an I-EDI system, benefit when new participants enter the platform (Katz and Shapiro, 1994). For example, the system becomes more valuable to any company as more of its suppliers and partners adopt the system, because that allows the company to exchange the information and data with them without any additional integration or implementation costs. This creates a positive feedback loop (Katz and Shapiro, 1994). When large corporations adopted the I-EDIs, their supply networks had to comply and adopt compatible systems themselves (Bednarz, 2004), which then encouraged other corporations to adopt them as well (Zhu et al., 2006) — such dynamics are one of key defining attributes of network effects (Katz and Shapiro, 1994; Rochet and Jean, 2003).

Such characteristics of network effects were, however, already present in the VAN-based EDIs discussed in section 2.2.1 and Takac (1992) already

[6]Transmission Control Protocol/Internet Protocol

[7]eXtensible Markup Language

Item	EDI	Internet-Based IOS
Data standards	Various	Open standards (XML-based)
Complexity	High	Low
Customization	Highly parter-specific	Less partner-specific
Communication protocols	VAN (private)	Internet (open, TCP/IP based)
Interoperability	Low	High
Communication costs	High	Low
Scope	Relatively narrow, with existing partners	Broad, with existing and new partners, hence strong network effects

Table 3: Comparison between EDI and internet based IOS (adapted from Zhu et al. 2006)

recognized that ”biggest savings will come only when and if the maximum number of companies—which will often be in competition—are connected through the same compatible networks”. The open-standards of the web, combined with low adoption costs, reduced the friction[8] of entering the network, allowing a large number of smaller companies to adopt interorganizational information systems as well (Stefansson, 2002). As all participants increase the net value of the network (Parker et al., 2016b), the network became increasingly attractive when more companies were able to adopt the systems. Hence, the network effects for I-EDI adoption and the subsequent value to adopters were accelerated significantly to levels that could not have been achieved with the VAN based EDIs.

Zhu et al. (2006) conducted a survey to study the migration to open-standard IOS and found network effects and expected benefits to be significant drivers of migration to open-standard IOS, and that companies that used prior EDI systems were reluctant to adopt and implement the new I-EDI due to the switching costs. It’s also an example of the pre-existing network effects hindering the transition to a new network (Katz and Shapiro, 1994) — the reluctant companies likely still used their EDI with their partners and sup-

[8] In platform business, the friction refers to the amount of resources any given party must commit to enter (Parker, Van Alstyne, and Choudary, 2016b)

pliers and without them transitioning first, the transition to a new network altogether would be costly and difficult. However the change had already become permanent as the large networks built around the internet based information exchanges had already started transforming the way information could be shared among organizations, resulting in radical transformation of organizational practices in procurement, deliveries, and financial transactions (Huang et al., 2008). Firms could now leverage the connectivity of the internet to enable real-time information sharing and improving coordination of allocated resources across the supply chain and with new partners altogether (Lee, 2004).

The internet allowed the data sharing to reach new levels altogether as the networks of linked parties grew — due to the low adoption costs, even small companies were able to leverage the benefits of EDIs and simultaneously strengthen the network as well. Additionally, the adoption of standardized data formats allowed companies to set up new data streams with much less effort, as additional integration was much simpler when their current systems were already able to interpret the messages. Moreover, the increased integration with third parties with these technologies—particularly in the supply chain network—created some strategic opportunities to the companies that were able to exploit the systems to their full potential (Auramo et al., 2005).

2.2.3 Application programming interfaces

Application programming interfaces (API) associated with various types of web services allow developers to easily integrate diverse content (e.g. apps) from different web-enabled systems (Chen, Storey, and Chaing, 2012). Jacobson et al. (2011) provide a following technical definition: "An API is a way for two computer applications to talk to each other over a network (predominantly the Internet) using a common language that they both understand". It is important to note, that the definition does not include anything about interorganizational information sharing — APIs are such a wide phenomenon, that they are used both internally and with other companies as well (Jacobson et al., 2011; Smith et al., 2016; Moilanen et al., 2019). Moilanen et al. (2019) also categorize some APIs as boundary resources which is largely the point of view from which this study addresses them. In addition to serving as an expanded means to data sharing, APIs may also facilitate the networking of disconnected pockets of expertise (Purvis, Sambamurthy, and Zmud, 2001), integration of new software into legacy software (Joseph, Ludford, and McAllister, 2016), speed IT deployment (Iyer and Subramaniam, 2015), and creation of entirely new products and offerings (Jacobson et al., 2011; Liozu and Ulaga, 2018; Moilanen et al., 2019). Technically, APIs are not widely different from the other internet based information systems as they use the same underlying technologies such as XML or JSON[9] over the web (Huang et al., 2008; Chen et al., 2012). However, due to the emergence of 'big data' and the increasing number of devices connected to the internet, it is considered as an evolution in the data sharing mechanisms.

APIs can also change the way companies do business — Moilanen et al. (2019) use the EU-wide banking legislation PSD2[10] as an example in which the banks must rethink their services and offerings when customer interfaces are no longer under their control.

Big data

Before addressing APIs in more detail, it is important to understand what big data means. Fosso Wamba, Akter, Edwards, Chopin, and Gnanzou (2015) conducted a review of the definitions of big data and concluded that big data is "a holistic approach to manage, process and analyze 5 Vs (i.e. volume, variety, velocity, veracity and value) in order to create actionable insights for sustained value delivery, measuring performance and establishing competitive advantages". Initially, the definition of big data was built on the first three Vs (McAfee and Brynjolfsson, 2012), but later literature has expanded the framework to include other attributes as well (Fosso Wamba et al., 2015; Günther, Rezazade Mehrizi, Huysman, and Feldberg, 2017; Tao, Qi, Liu, and Kusiak, 2018). Table 4 summarizes the descriptions of the attributes.

Leveraging the study by Chen et al. (2012), we can outline some of the drivers for the growth of big data. He argues that there has been three eras of business intelligence (i.e. the kind of data available from companies' operations) that are roughly tied into this same evolution; (1) at the time of EDI, business intelligence included primarily data of inventory levels, possibly some point-of-sales data, and other specific operational attributes (see e.g. Reekers 1994; Ahmad and Schroeder 2001). After the widespread adoption of the internet, the second (2) era of business intelligence allowed tracking users' web behavior through web analytics (Huang et al., 2008; Chen et al., 2012), which enabled much more holistic analysis of the users' and companies' patterns and preferences. The third era (3) is driven by the shift in

[9] JavaScript Object Notation

[10] Payments services (PSD 2) - Directive (EU) 2015/2366 will force all banks in the EU to open APIs for external parties and developers. In practice, it means that banks have to provide access to the core functions of the bank (i.e. making payments and deposits) for any application the customer has authorized. For more information: `https://ec.europa.eu/info/law/payment-services-psd-2-directive-eu-2015-2366_en` Accessed 18.06.2019

Attribute	**Description**
Volume	Large volume of data that either consumer huge storage or consist of large number of records
Variety	Data generated from greater variety of sources and formats, and contain multidimensional data fields
Velocity	Frequency of data generation and/or frequency of data delivery
Veracity	Inherent unpredictability of some data requires analysis of big data to gain reliable prediction
Value	The extent to which big data generates economically worthy insights and or benefits through extraction and transformation

Table 4: Attributes defining big data (adapted from Fosso Wamba et al. 2015)

internet usage — The Economist (2011) reported that in 2011 the number of mobile devices connected to the internet surpassed the amount of PCs. Such personal devices that are used constantly also gather location data from the users, and the analysis can therefore be context aware and more personalized. Considering the attributes of big data (Fosso Wamba et al., 2015) in table 4, the third era of business intelligence fits the framework really well; volumes are massive and variety is ensured due to the amount of different devices connected to the internet (Manyika, Chui, Brown, Bughin, Dobbs, Roxburgh, and Hung Byers, 2011), velocity is guaranteed due to the constant connectivity of the devices, the usage of mobile devices for all aspects of life ensures the veracity of the collected data, and finally several companies have been able to leverage significant value from location data alone through analytics (Manyika et al., 2011). For more discussion about the value derived from big data, refer to section 2.3.1.

APIs are one way of making the big data available to several parties as it allows companies and users to distribute and read data from various sources with calls over the internet, allowing the development of various applications (Jacobson et al., 2011). APIs are code that govern the type and format of calls that any given application can make of another associated program (Benzell, LaGarda, and Van Alstyne, 2017). The associated program is agnostic about the source of the call and the app does not need to know anything about the internal workings of the associated programs — it merely answers the call.

APIs in business

As discussed previously, APIs do not necessarily differ from the interfaces provided by open standard interorganizational information systems operating over the internet from a technical point of view (Huang et al., 2008). The key difference is the emergence of "open APIs"[11], also Web APIs or public APIs, which are interfaces from which any party or application can make calls from over the internet (e.g. Jacobson et al. 2011). Programmable Web, an internet repository of Web APIs has recorded the number of open APIs since the first public API in 2005. In their latest release, they have over 17000 APIs on record (Santos, 2017). Hence, there are massive amounts of data available openly to any company or an individual, which makes the networks significantly larger than in the era of I-EDIs. The value of such networks, using the same logic as in section 2.2.2, would then be significantly larger as well (Parker et al., 2016b). However, we must consider how the network effects are formed; now there are countless entities making API calls to extract data (Benzell et al., 2017), but without them adding any value to the data providers, the network effects don't necessarily become stronger (Rochet and Jean, 2003; Parker et al., 2016b). On the network effects of API economy, Moilanen et al. (2019) argue that developers are drawn to API providers when there are a range of solutions already available to speed up and facilitate the development of new services which further emphasizes the fact that network effects are only created around the best solutions.

From a business point of view, switching to API based service development allows companies to radically reduce development time, shortening the time-to-market for solutions — instead of a traditional integration project over several months, business collaboration with companies can be done using APIs in weeks or even in a matter of days (Moilanen et al., 2019). Smith et al. (2016) identified five tangible benefits of APIs that create value for companies:

1. The API reduces the complexity experienced by standardizing the implementation of the tasks

[11] However, the bulk of APIs are still private (Jacobson et al., 2011) and used either within organizations or as proprietary means of information exchange between companies. Such proprietary means can be e.g. paid access to a real-time feed of a specific set of data, but these are not technically different from the more advanced I-EDIs.

2. The API provides better access to information or data by providing an open (or limited, depending on desired choices) availability

3. The API increases the chances of influencing content development by bringing new channel for data providers and, on the other hand, users' needs are reflected in API's demand

4. The API reduces the perceived risk because the dependence on a single API is small, and on the other hand the API is interchangeable with another with a relatively minor effort

5. The API enhances the visibility of the service, encourages open innovation, and illustrates the benefits of an open data policy.

Looking back at the benefits of EDI and I-EDI, the evolution to API is clear. The reduced complexity has evolved in each step through standardization, and the access to the data has become much easier over time. APIs, however, have also been able to bring forth new benefits altogether as we can see from the list above (Smith et al., 2016).

Emergence of machine-to-machine communications

Nowadays data and information sharing is increasingly taking form as machine-to-machine (M2M) communications. In their book, Brynjolfsson and McAfee (2014) already outline how M2M communications take place in real time, all the time, with data being exchanged across various boundaries without any human interaction. This has been enabled by mobile internet connectivity — prior to mobile broadband connectivity such communications were not feasible (Wu, Talwar, Johnsson, Himayat, and Johnson, 2011). M2M communications also use other connectivity methods, such as an aggregating gateway combined with lower-cost radio protocols that forms a small proprietary networks of machines that communicate with each other (Wu et al., 2011). Companies can, for example, embed sensors and radios to their machinery that dynamically adjust the flow of the production based on the data gathered in any given step in the process as the machines react to each others' current statuses. Wu et al. (2011) anticipate future M2M ecosystems to face explosive growth—given successful standardization—across industries and become very complex in the process.

In addition to data throughput and connectivity, one of the key concerns in M2M development is the energy efficiency of the modules (Wu et al., 2011). Upcoming fifth generation (5G) of cellular connectivity is expected to play a critical role in the growth of M2M communications due to its low latency, energy efficiency, and high data capacity (Schulz, Matthe, Klessig, Simsek, Fettweis, Ansari, Ashraf, Almeroth, Voigt, Riedel, Puschmann, Mitschele-Thiel, Muller, Elste, and Windisch, 2017).

Looking at the evolution of data sharing mechanisms and the underlying technologies, from EDIs to APIs, three overarching areas have drastically evolved in parallel to the technologies; (1) the scope and amount of data that can be shared both within companies and with external partners has expanded from purely operational data to all-encompassing big data, (2) the expected benefits of data sharing initiatives have evolved from specific operational efficiencies to enabling new strategic opportunities, and (3) the number of stakeholders involved in the data sharing networks has grown from few individual partners to countless potential partners as the means of communication have been standardized. This is also in line with the study by World Economic Forum[12] in which the device-level technologies from mainframes to mobile are interlinked into the disruptive technologies.

2.3 Role of data on firm performance

A key attribute of information and data, is that it is non-rival in nature — it can be copied and reused infinitely and it is not consumed when used (Brynjolfsson, Hitt, and Kim, 2011; Brynjolfsson and McAfee, 2014). Many companies tend to overlook the immense potential of accumulated historical data, because replicating the data is impossible and approximating it becomes increasingly difficult the more extensive the data is (Liozu and Ulaga, 2018), which underlines the value and importance of having first hand access to such data. In regards to sharing, benefits of data sharing specifically can be derived from the reduction of information asymmetries[13]; once all information is available to all economic parties, there are no inefficiencies for intermediaries to take advantage of, and new solutions can be developed that harness all of the available information. In this section, the various benefits of data are assessed through operational and strategic lenses. There is extensive literature addressing the benefits of information and data sharing in operational context (e.g. Dawes 1996; Gavirneni et al. 1999; Lee et al. 2000), and some literature on the benefits of exploiting big data (e.g. McAfee and Brynjolfsson 2012; Mayer-Schönberger and Cukier 2013; Fosso Wamba et al.

[12]`https://www.weforum.org/agenda/2019/05/we-need-to-measure-innovation-better-heres-how-to-do-it/` Accessed 18.06.2019

[13]Information asymmetries have been studied across scientific fields, for implications of information asymmetries on financial markets see e.g. Chung, McInish, Wood, and Wyhowski 1995

2015) — because the approach of this study has not been taken before, the insights must be derived by combining these approaches.

2.3.1 Operational efficiencies

As discussed in section 2.2, the most notable benefits enabled by more sophisticated data sharing systems were initially operational. EDIs allowed much more efficient information sharing between organizations and their suppliers which resulted in cost reductions associated with handling the information and cutting inefficiencies, such as too large inventories, driven by the information asymmetry (see e.g. Reekers 1994; Mukhopadhyay et al. 1995). However, the operational efficiencies enabled by information sharing are a wider phenomenon than the benefits of adopting information sharing systems. In this section, the impact of data and its sharing on current business models will be reviewed in detail.

Data-driven decision making

Business intelligence and analytics (BI&A) has emerged as an important area of study, which reflects the magnitude and impact of data-related problems to be solved in contemporary business organizations (Chen et al., 2012). While business intelligence has been associated with data since the 1990s, business analytics was introduced in the late 2000s to represent the key analytical component in business intelligence (Davenport, 2006; Chen et al., 2012). Davenport and Kudyba (2016) state that data and analytics were initially named as "decision support" aimed to improve the accuracy and efficacy of decisions, but the emergence of big data has enabled more varied use of business analytics. A report by McKinsey Global Institute estimates \$600 billion potential annual consumer surplus from using personal location data globally and 60% potential increase in retailers' operating margins made possible with analytics and big data (Manyika et al., 2011). Although the numbers are difficult to verify and should not be taken at face value, they do illustrate the business potential unlocked by having access and deriving analysis from big data. A substantial amount of company, industry, product, and customer data can be gathered from the web using cookies and server logs — by analyzing the data with tools such as Google Analytics, companies are able to determine the user's browsing and purchasing patterns (Chen et al., 2012). However, it is important to emphasize that the value is realized only when big data is used for something as it does not have any value in its own right (Liozu and Ulaga, 2018), which is why BI&A is such an important area for any business. When utilizing the data correctly, the results can be very attractive; McAfee and Brynjolfsson (2012) concluded in their study that data-driven businesses perform better than their competitors with an average of 6% more productivity and 5% more profitability.

Increased visibility of information and data within organizations can also lead to improved productivity — when information is shared across various boundaries within companies or their networks, it's possible to discover patterns and interactions that would have been indistinguishable otherwise, and to use these discoveries to make data-driven decisions (Dawes, 1996). Additionally, having access to big data, through either trading or sharing, enables predictive analytics (Davenport and Kudyba, 2016) which allows companies to do more accurate forecasts that are relevant to their business. One of the most common forms of predictive analytics is predictive maintenance for industrial machines — using data gathered from sensors in machines, either the company's own or comparable machines from other companies, analytics compute the point when comparable machines have broken down and recommend particular services before that time (Davenport and Kudyba, 2016) and such analytics are becoming a key enabler for enhancing manufacturing competitiveness (Tao et al., 2018) as it allows for superior efficiency over legacy models.

Supply chain efficiencies

The effects of information sharing in supply chain has been extensively reviewed in the literature (see e.g. Stevens 1989; Gavirneni et al. 1999; Cachon and Fisher 2000; Lee et al. 2000; Lee and Whang 2000; Yu, Yan, and Cheng 2001; Lau et al. 2003; Vickery, Jayaram, Droge, and Calantone 2003; Barratt and Oke 2007; Prajogo and Olhager 2012; Lotfi, Mukhtar, Sahran, and Zadeh 2013) and the conclusions remain largely unified; information sharing improves efficiency in the supply chain. Studies on information sharing technologies, such as EDI, can be used to identify the specific mechanisms of information sharing that can be improved with technology, but the underlying benefits are unlocked due to more efficient information sharing.

Data sharing with suppliers and partners can be considered as one form of supply chain integration (Vickery et al., 2003) and its value can be traced to the Value Chain Model developed by Porter (1980, 1985). In the value chain model, "linkages" are the relationships between the way in which one value activity is performed and the cost or performance of another. Optimizing these linkages is the core purpose of supply chain integration, and literature has found evidence that such integration should engender superior performance (e.g. Tan et al. 1998; Westbrook and Frohlich 2001). In more recent literature, Baihaqi and Sohal (2013) already consider close coordination and information sharing in supply chains to be necessary in order to maintain competitive advantage in the more demanding markets.

Literature also shows how the lack of integration or information sharing can be harmful for the company's performance. The 'bullwhip effect', a cause of unnecessary wastes and a term used to determine a higher amplification of order and inventory fluctuations is driven by insufficient information exchange between parties in the supply chain (Lee, Padmanabhan, and Whang, 1997; Yu et al., 2001; Lau et al., 2003). The main source for the amplification and fluctuation is driven by the lack of timely sharing of production information between enterprises in the supply chain (e.g. Lee et al. 1997) and the subsequent uncertainties are usually buffered by inventories (Yu et al., 2001). Hence, the benefit of information sharing is significant in reducing the bullwhip effect and supply chain costs when each party can make better decisions on ordering, capacity allocation and production/material planning (Lee et al., 1997, 2000; Lau et al., 2003). Additionally, in their modelling of the value of information, Gavirneni et al. (1999) found out that information is always beneficial when assessing costs and savings due to having to hold less inventory as a buffer for uncertainty.

To illustrate what kind of data companies can gather and utilize, Najjar and Kettinger (2013) identified three types of data retail companies gather: point-of-sales data, customer loyalty data, and inventory data. Specifically point-of-sales data and inventory data are such that enable the aforementioned supply chain efficiency benefits (Cachon and Fisher, 2000; Najjar and Kettinger, 2013) and analytics on customer loyalty data can be used for various strategic opportunities.

Driver	Selected references
Smaller inventories	Mukhopadhyay et al. (1995); Lee et al. (2000); Lee and Whang (2000); Yu et al. (2001); Lau et al. (2003)
Quicker lead times	Lee and Whang (2000); Cachon and Fisher (2000)
Reduced payment cycles	Cachon and Fisher (2000); Lau et al. (2003)
Reduced labor costs	Lee and Whang (2000)
Mitigation of the bullwhip effect	Lee et al. (1997); Lee and Whang (2000); Yu et al. (2001); Lau et al. (2003)

Table 5: Supply chain efficiency drivers

Table 5 summarizes the levers through which data-related supply chain efficiencies are realized. Most of the efficiencies are due to reduced inefficiencies between the supply chain parties which are reliant on sharing the relevant data and information with them.

2.3.2 Strategic opportunities

In addition to improving current operations through factors such as transparency and reducing inefficiencies, the rapid growth of available data has also created new strategic opportunities for old and new companies alike. New innovations can be developed (Constantiou and Kallinikos, 2015) and business models that have never existed before have been introduced, and new companies have emerged that have seized the opportunity to create business on trading data or acting as data platforms. For example, credit card companies have recognized the value of collecting and aggregating the data collected from the consumer spending (Nunes and Drèze, 2006) and selling the consumer behavior data to businesses. As another strategic implication, Liozu and Ulaga (2018) argue that companies' current customer segmentation will differ from data-enabled customers — indicating that partaking in data sharing and trading allows companies to make business with new customers altogether.

Augmenting current business and developing new offerings

Since big data can significantly affect companies' performance (e.g. McAfee and Brynjolfsson 2012; Mayer-Schönberger and Cukier 2013; Fosso Wamba et al. 2015), some companies have seen the opportunity to develop new products and services leveraging the opportunity. For example, in the agricultural products industry, there are companies that offer data products that recommend when and what farmers should plant, when the plants should be watered, and when to harvest them (Bunge, 2014). This is an example of prescriptive analytics, which recommends specific actions based on historic and real-time data on similar situations — other forms of prescriptive analytics involve algorithms which match customers with products, with dating candidates, or with potential business network members (Davenport and Kudyba, 2016). Such analytics are fuelled by data, and companies who gather this data can exploit such opportunities to create new business by either exploiting the advantage themselves, or by providing such analytics as services to other companies (Najjar and Kettinger, 2013). Similarly, the predictive analytics used for operational efficiency in manufacturing business could be exploited further by creating new service business models, in which a machinery company could start selling the analytics from their machines to other manufacturers (Liozu and Ulaga, 2018; Tao et al., 2018). In this case, the seller would benefit from the monetary value of the analytics, and the buyer could realize the operational efficiencies the larger company had found through optimizing the larger installed base of machines.

Data monetization can be defined as "the act of exchanging information-based products and services for legal tender or something of perceived equivalent value" and they can include a range of offerings from raw data to process design (Woerner and Wixom, 2015). In the same paper, the authors argue that data monetization is done by wrapping, selling and bartering — wrapping information around other core products and services, receiving money in exchange for information offerings, or by choosing to trade information in return for new tools, services, or special deals. For further verification, Wixom and Boss (2017) also identify both wrapping and selling data as the two main ways of data monetization but they propose improvement of internal business processes and decisions—which is considered as an operational benefit in this thesis—as the third way of data monetization. However, raw data alone has no value without any insights derived from it[14] (Liozu and Ulaga, 2018) but it can immensely valuable to an actor who is capable of conducting the correct analysis on the data.

Predictive analytics discussed in section 2.3.1 can also be utilized as a

[14] For more discussion on insights derived from data, see section 2.3.1

powerful tool in strategy making, as strategy is based on making predictions of future events based on collected data and information (Constantiou and Kallinikos, 2015) — big data about the whole internal and external business environment can be utilized to make data-driven insights in real time (Constantiou and Kallinikos, 2015; Woerner and Wixom, 2015) making e.g. reaction times to market trends significantly lower. Therefore it is justifiable to argue that data sharing and analytics can also improve companies' strategies by improved decision making. Woerner and Wixom (2015) identified three distinct ways in which big data supports the measurement and monitoring of strategy; (1) new data altogether from all areas of the business, (2) new insight from various visualization and business intelligence tools, and (3) new actions that are enabled by the real-time metrics and trends. They also note that companies do not have to replace their existing toolboxes in order to leverage the opportunity, but rather they can use their existing toolboxes more effectively as they now can have access to data needed to solve problems or gain insights (Woerner and Wixom, 2015).

New market opportunities and stakeholders

The strategic opportunities enabled by more open data sharing can also extend beyond the company borders and its current market, as the new products and services can create new markets altogether. However, there is very little literature on this specific phenomenon but there are some examples in which external parties have created entirely new markets due to the practices of a specific firm.

The strategic opportunities enabled by APIs and big data are still being discovered — for example in the context of the PSD2 directive, some banks are considering monetizing access to "premium APIs" which would create an entirely new source of revenue from the application developers (Moilanen et al., 2019). Another strategic option for any business is taking advantage of the potential value of data and monetize it.

The notion of complementary assets traces to Teece, Pisano, and Shuen (1997) and it has been refined into complementary innovations—in which a third party develops a product or a service that relies in an existing product or service—in information technology studies (Brynjolfsson and Hitt, 2000) and in platform economy as well (Parker et al., 2016b). Apple's App Store is a key example; there are almost two million[15] applications that are developed on Apple's iOS platform and they generate billions of dollars[16] to

[15]`https://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/` Accessed 18.06.2019

[16]Apple does not report the exact App Store revenues, but their services revenue was

Apple annually, and that is enabled by Apple freely sharing the necessary software development kits and documentation to the application developers and allowing them access to the iOS platform (Parker et al., 2016b). As Tiwana et al. (2010) note, such software development practices allows the focal company to access skills and an appreciation of user needs that the platform owner does not internally possess — allowing the development of entirely new capabilities and applications unforeseeable by the platform's original designers.

Another way to create new business from data sharing is by allowing external parties access to your data, and make use of mass collaboration (Tapscott and Williams, 2006). Giving external parties access to data allows third parties to develop their own business as complementers much like in the App Store example (Parker, Van Alstyne, and Jiang, 2016a; Parker et al., 2016b; Zhu and Liu, 2018; Parker and Van Alstyne, 2018). In practice, it already creates a simple ecosystem[17] as the businesses require the platform or data to stay in business, and the ecosystem owner ideally benefits directly from the additions. While many of the examples used in platform literature revolve around the information technology industry, other industries have been able to leverage complementers as well — a gold producer, who opened its geographical databases to the public and offered prizes to those who were able to tell where they would find gold, was able to increase its gold production by 850% and cut its production cost per ounce by 84% in just one year (Peppard, Edwards, and Lambert, 2011). This is an another example in which the true innovation is happening beyond the company's own business and the results benefit both the incumbent company as well as the third parties developing the solutions. Such dynamics and opportunities beyond companies' borders have not yet been widely researched but it is important to recognize how data sharing can play a crucial role in such opportunities.

2.3.3 Dimensions of data benefits

Tying together the benefits discussed in the previous sections, a following categorization is proposed in table 6.

approximately $37 billion in 2018

[17] In a business context, an ecosystem is a network of interacting organizations from various functions involved in (co-)creating goods and services, and the actors of the ecosystem typically both compete and cooperate with each other (Liozu and Ulaga, 2018)

Dimension	Sub-dimension	Definition
Strategic opportunities	Internal	In-house initiatives such as new products and services driven e.g. by access to big data and analytics
	External	Opportunities developed beyond company boundaries but which directly benefit the observed company
Operational efficiencies	Internal	Efficiencies realized by improving current operations e.g. reducing labor costs
	External	Benefits allowing external stakeholders to operate more efficiently, leading to reduced costs for the observed company

Table 6: Benefits of data

The dimensions of the framework tie into the sections 2.3.1 (operational efficiencies) and 2.3.2 (strategic opportunities) and the sub-dimensions are derived from the contents of those sections. The most common and straightforward benefits are definitely in the internal dimensions and the extent of data sharing practices defines whether the strategic opportunities are unlocked or not; internal operational efficiencies can be realized as the increased visibility within processes eliminates some of the need for labor (e.g. Barratt and Oke 2007) but strategic opportunities such as product innovations likely require more open data sharing, or access to big data as proposed by e.g. Woerner and Wixom (2015) and Constantiou and Kallinikos (2015), to incubate new ideas.

It is important to highlight that the internal dimensions in table 6 can be realized by buying the data from other parties or collecting it internally, and then applying analytics tools on it. The external benefits, however, can only be realized if data is shared with external parties soince, by definition, the efficiencies or innovation is happening outside the company borders.

In this chapter, the evolution of data sharing mechanisms has been thoroughly outlined to understand how APIs can be considered as a natural evolution of EDI systems. Simultaneously, the benefits of data discussed in literature have largely been dependent on the technological era — only operational efficiencies were addressed in 1990s and the strategic value of data has been reviewed in approximately the past decade. Formed after the extensive literature review, a framework for categorizing these benefits was introduced in the end of this chapter. At this stage, these results remain theoretical and conceptual and must be tested with empirical data.

Chapter 3

Methodology

Since this study can be described as a theory building effort, the empirical research will be focused on validating and refining the theoretical framework. The chosen methodology in this study is triangulation[1], in which multiple methods and data sources are used to enhance the validity of the research findings (Mathison, 1988; Greene, Caracelli, and Graham, 1989). Triangulation methods have been utilized in various fields of research, and Lewis (1998) provides a comprehensive overview of how it is applicable in operations management research and Jonsen and Jehn (2009) apply the methodology to management research. It is well applicable to this study since its purpose can be defined as follows: ”... triangulation is supposed to support a finding by showing that independent measures of it agree with it or, at least, don’t contradict it” (Miles and Huberman, 1984). The need for triangulation is therefore clear — because no previous research from the same perspective exists, the techno-economics of data sharing must be understood from different perspectives in order to ensure the rigidity of the framework, and to form a holistic[2] understanding of the dynamics. Hence, both quantitative and qualitative research methods will be used to further develop the theoretical framework. Combining quantitative and qualitative methods is also the most commonly used type of triangulation in business and management literature (Jick, 1979; Modell, 2005), which supports the validity of this approach. Additionally, an abductive approach is applied to refine the theory. Dubois and Gadde (2002) suggest systematic combining as an approach for developing theories from case research — such an approach fits well with tri-

[1]Triangulation traces back to Campbell and Fiske (1959) where the authors argued that more than one method should be used in the validation process to ensure that the variance reflected that of the trait and not of the method

[2]Jick (1979) argues that triangulation can capture a more complete, holistic, and contextual portrayal than using a single type of research method.

angulation and to the context of this study as well, as the theory is developed further with the help of experts (i.e. cases).

3.1 Theoretical approach

In chapter 2, the benefits of data sharing were linked with the advances in technology. Based on these findings, a theoretical framework can be developed as seen in figure 3. The sub-dimensions of benefits (internal and external) are left outside the figure for the sake of clarity.

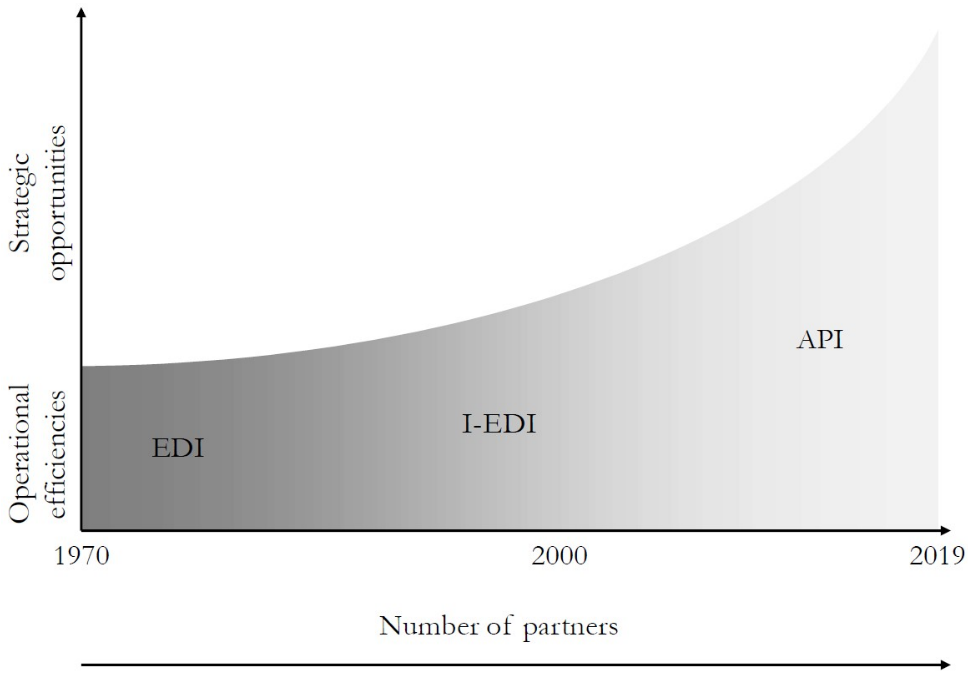

Figure 3: Proposed technological framework

Based on the literature, 3 hypotheses are formed for the empirical research to validate the assumptions of the proposed framework.

> **Hypothesis 1:** Historically, the penetration and adoption rates of sophisticated data sharing technologies has grown with technological advancements

The literature extensively talks about the decreased adoption costs and increased benefits of more sophisticated data sharing systems (e.g. from EDI to I-EDI) which should result in increased adoption and subsequent readiness for data sharing systems (e.g. Bednarz 2004; Zhu et al. 2006; Narayanan et al. 2009).

> **Hypothesis 2:** The evolution of data sharing technologies has allowed companies to realize more benefits from data sharing practices.

When EDI was the dominant data sharing technology, the majority of recorded benefits from data sharing were operational efficiencies, specifically from supply chain management (e.g. Lee et al. 2000; Lee and Whang 2000). However, the emergence of more advanced analytics and availability of big data has created new business opportunities altogether (e.g. Peppard et al. 2011; Chen et al. 2012; Constantiou and Kallinikos 2015). Therefore I hypothesise that the evolution of data sharing technologies—and the subsequent increased variety of shared data—has been linked to these benefits during the past decades.

> **Hypothesis 3:** Data enabled benefits can be categorized into operational efficiencies and strategic opportunities, with further division into internal and external benefits.

The third hypothesis is formed to validate the lenses for addressing benefits enabled by data on page 30.

3.2 Data collection and preparation

To conduct sufficient triangulation, the empirical research of this study consists of two complementing parts; quantitative historical analysis of company surveys will be used to validate the assumptions about data sharing system adoption and several expert interviews are conducted to gather empirical evidence of the evolution and the benefits of data sharing. Additionally, I had informal discussions with a CEO of a data services company in order to understand the business logic more thoroughly.

3.2.1 Quantitative analysis

In order to research the adoption of technologies linked to data sharing, I used the data from Statistics Finland, specifically the annual data of "ICT and e-commerce in companies" from years 2003–2016 and bi-annual "Innovation research" from years 2002–2016. The data sets are gathered from mandatory surveys across all companies in Finland and there are approximately 3000 responses each year in the ICT survey and approximately 2500 responses in the innovation survey. The data from different years were consolidated into a single database in which annual percentages were calculated for each variable as seen in equation 3.1 and only companies with more than 10 employees were included.

$$\frac{1}{n}\sum_{i=1}^{n} x_i, x \in \{0,1\} \tag{3.1}$$

In equation 3.1 n denotes the number of respondents in each survey and x denotes the responses i.e. 1 = yes and 0 = no. This results in share of companies that responded positively to any given question in a single survey. However, extensive trends over the entire time period are impossible to conduct due to many of the questions changing through the years.

Additionally, I was able to use the results of "The importance of data in business" -survey conducted by The Research Institute of Finland in 2015 in which there are responses from approximately 1300 Finnish companies across all industries.

ICT and e-commerce

The annual survey data of ICT and e-commerce of Finnish companies by Statistics Finland is used to determine the adoption rate of ICT technologies that are associated with data sharing. Unfortunately, the survey does not have specific answers regarding data sharing or data sharing technologies, but by assessing several related variables (as seen in table 7) the high level trends can be observed instead. However, only trends regarding the EDI and I-EDI can be observed from this data, as there is no data regarding API usage or historical details about exploitation of big data and analytics. Additionally, the compounded annual growth rates for the various variables were manually calculated.

Variable & available years	Description	Reasoning
edi *(2003–2007)*	The firm uses EDI	Very straightforward question to address EDI penetration
edivast *(2003–2007, 2009–2016)*	Firm received EDI orders	Gives and indication of EDI usage and is available over a long time period
autedi *(2008–2012)*	EDI standards or equivalent in automatic data exchange	Indication of EDI usage, comparable to autxml
autxml *(2008–2012)*	XML based standards in automatic data exchange	Comparable to autedi to see transition to open standard systems
ostedi *(2009–2015)*	Firm makes purchases using EDI	Comparison with ostweb
ostweb *(2009–2015)*	Firm makes purchases using web	Comparison with ostedi
intra *(2003–2010)*	Firm has an intranet	Comparison with extranet
extra *(2003–2010)*	Firm has an extranet	Data sharing system with external parties
scm *2007–2010, 2012, 2014–2015*	Company shared data electronically in its supply chain	Indicator of operational data sharing penetration

Table 7: Variables selected from the ICT and e-commerce survey

Innovation research

The innovation research from Statistics Finland is conducted bi-annually and focuses on various innovation practices conducted by companies operating in Finland. It was selected to this study because the survey addresses cooperation with external parties in innovation processes which is interpreted as an indicator of readiness in data sharing practices since companies cannot cooperate with no data sharing between them.

If the responding company had innovation activities together with another entity, the survey had different variables to specify the geography of the partner. Grouping variables were created in order to assess whether the respondents were cooperating with various entities at all, regardless of the geography (e.g. *cosup* = 1 if the company has done innovation cooperation with suppliers either in Finland, EU, outside EU, or elsewhere).

Additionally, the latest revisions of the survey have questions about exploiting data in the companies which can be used to further develop understanding of how large of a role does data play in companies nowadays.

Variable & available years	**Question**	**Reasoning**
rrdin *(2000, 2004, 2008–2016)*	The company has done product or service innovations internally	Indicators of how much innovation is done internally versus externally
rrdex *(2000, 2004, 2008–2016)*	The company has done product or service innovations through external partners	
dataosto *(2014, 2016)*	The company has bought big data to other companies	Used to understand the value of data and the willingness to share or trade it
datamyynti *(2014, 2016)*	The company has sold big data to other companies	
dataprosein *(2014, 2016)*	Data is used in process innovation	Understanding the important use cases for data
dataorgin *(2014, 2016)*	Data is used in organizational innovation	
datamarkin *(2014, 2016)*	Data is used in marketing innovation	

datatk *(2014, 2016)*	Data is used in research and development processes	
datamarkk *(2014, 2016)*	Data is used in marketing	
co *(2000, 2004–2016)*	Company does cooperative innovation	Understanding the readiness to share information through cooperation
cogrp* *(2000, 2004–2016)*	Company does cooperative innovation within its own corporation	
cosup* *(2000, 2004–2016)*	Cooperative innovation with suppliers	
cocli* *(2000, 2004–2016)*	Cooperative innovation with clients	
cocom* *(2000, 2004–2016)*	Cooperative innovation with competition	
cocon* *(2000, 2004–2016)*	Cooperative innovation with external consultants or private research centers	
couni* *(2000, 2004–2016)*	Cooperative innovation with universities	
cogmt* *(2000, 2004–2016)*	Cooperative innovation with public research centers	

* custom grouping variable

Table 8: Variables selected from the innovation research survey

The shares of respondents were calculated for each year and the compounded annual growth rates were manually calculated over the available time periods using the following equation.

$$\left(\frac{a}{b}\right)^{\frac{1}{c}} - 1 \tag{3.2}$$

In equation 3.2 a denotes the latest available value, b denotes the earliest available value, and c denotes the number of years between the observations. The resulting compounded annual growth rate can be then used to compare the growth of various variables regardless of the differences in observed time frames.

The importance of data in business

In 2015, The Research Institute of Finland conducted a survey regarding the use of data in Finnish companies and there were 1303 respondents in total. The survey focused on the use of big data but also included questions regarding the sharing and trading practices of big data. In addition to the data questions, there is a question whether an open API is connected to any of the company's products which gives a numerical estimate on the adoption in the right-end of the framework.

Although the sample is smaller and there is no historical data available regarding these questions, the exact nature of the questions allows the data to be used to inspect the current adoption of APIs, use of big data, and the data sharing practices.

The questions addressing big data sharing and trading were responded to through sliders in which the respondent had to choose a value between 1 and 16 that reflected the company's practice of the given activity (1 = not at all, 16 = very much). For simplification, the responses were categorized into four categories based on their values:

- 1–4 Not at all or very little
- 5–8 Some but not much
- 9–12 Somewhat regularly
- 13–16 Quite much or a lot

With such categorization, it is easier to highlight the extent to which the companies share or trade big data with external organizations. The other values in the survey responses were either percentages or binary yes / no questions assessed similarly to the ones from Statistics Finland.

3.2.2 Expert interviews

Due to the limited availability of relevant data, 6 experts were also interviewed in order to qualitatively understand the whole evolution of data sharing mechanisms and the related changes in data sharing practices and benefits and validate the applicability of the theory developed in this study. The have each worked in companies that have gone through the various stages of the proposed framework and have experience in the impact of data on firm performance. The interviews were all conducted in Finnish, so the quotes are translations made by the author of the study, and as such are not exact quotations in regards to wording.

The reference interview guide can be found in appendix A but the interviews did not follow the guide exactly, rather they were arranged as semi-structured and divided into three sections: (1) to discuss the interviewees experience of the evolution of data sharing mechanisms, their benefits, the extent of their use, and how has the adoption of various technologies changed the company's data sharing practices; (2) the interviewees experience of data sharing and trading and its impact on companies' performance; (3) presentation of the theoretical framework, discussion of its validity and applicability, and the implications of the theoretical model to companies — the interview guide was used to steer conversation to the relevant topics if necessary. Hence, the interviews can also be considered thematic interviews as discussed by Aronson (1995). Each session was recorded and two hours were reserved for each interview and the short descriptions of interviewees can be found in table 9. Before conducting the interviews with the industry experts, a semi-structured interview was conducted with a professor in Aalto university who is familiar with the topic in order to validate the interview guide and gain some preliminary thoughts on the frameworks from academia.

Position and expertise	Date	Duration
Professor, previous career in digital financial services development	17.05.2019	1 h 25 min 17 sec
Retired CEO and board professional in credit card and financial sector	28.05.2019	1 h 37 min 50 sec
Senior advisor, ex-CEO in a large bank and current chairman in four companies	29.05.2019	49 min 50 sec
IT Architect in Data Management, data management experience from various companies and industries	03.06.2019	1 h 34 min 55 sec
Head of Ecosystems Research, decades of experience in standardization	04.06.2019	1 h 18 min 31 sec
CEO of an international payment services company	19.06.2019	1 h 15 min 47 sec

Table 9: Conducted expert interviews

After the first interview, the professor was positive about the validity of the proposed framework. The interview questions were tweaked according to his feedback to focus more on the areas in which relevant insights would be likely found.

Once all expert interviews had been conducted and recorded, transcripts were manually made from the recordings. The interviewees' comments were roughly categorized under the hypotheses of this study, and the rest formed thematic groups. The theoretical framework and findings were then reflected to the experts' comments and experiences in order to find contradictions and support for the hypotheses.

3.3 Data analysis

As is common in organizational research (Jick, 1979), the qualitative research was used to complement the quantitative results. The data from the can only give high level observations regarding the factors such as penetration, but there is no way to quantitatively address the perceived benefits, reasons for companies' adoption rates, or the hypothesised link to strategic benefits. Hence, the expert interviews are used to understand the context of the data, to find the reasons for the trends, and to complement the empirical data with the companies' point of view.

The statistical data was first gathered into tables and the trends were observed through the annual changes. Relevant calculations were made from each dataset and preliminary findings were noted down. The questionnaire was tweaked according to the data to ensure a holistic view.

Notes were made and taken from each interview, and the insights from the experts were consolidated under the main categories i.e. the phases of the technological evolution, and to the benefits enabled by data. All of the experts' observations were then reflected to the literature to find possible contradictions, complement the devised framework, and to validate the hypotheses. The analysis process was very qualitative in nature, and it was conducted in collaboration with a professor familiar in the subject.

Chapter 4

Analysis

In this chapter, the empirical data will be analyzed to validate the theoretical framework proposed in chapter 3. This chapter is divided into two sections addressing the importance and adoption of data sharing technologies, and the impact of data on firm performance individually.

4.1 Evolution of data sharing technologies

Overall, the interviewed experts were not able to provide detailed answers regarding various technology solutions due to the events taking place decades ago and them primarily focusing on the implications rather than the solutions themselves. However, with some probing, some practices and phenomena could be linked to the technologies based on the time frame or other relevant details in their experiences.

Interestingly, one interviewee wanted to highlight the importance of various events in information sharing context — before internet or other technology solutions, executives from various largest companies in Finland tended to meet up for dinner or sauna, and they would discuss various trends in the markets and in the business. Although not a data sharing technology, it highlights how common and important such information sharing has always been; information technology is merely a tool for facilitation rather than means to an end.

4.1.1 Pre-internet era

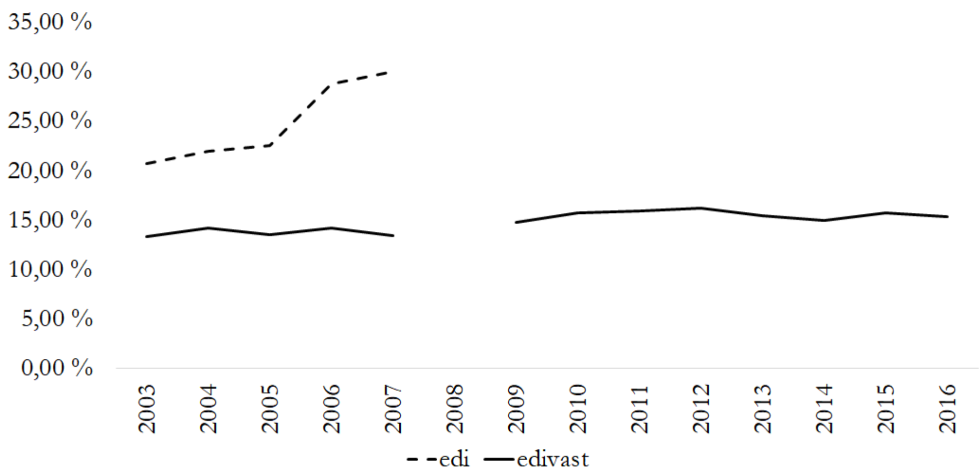

Figure 4: Share of companies using EDI or receiving orders via EDI

As seen from the figure 4, EDI adoption had largely matured and stabilized by the time the first data point of this survey was gathered. From 2003 to 2016 the CAGR of *edivast* is approximately 1.1% suggesting minimal growth — *edi*, on the other hand, from 2003 to 2007 grew with a CAGR of 9.7% which is curious. It may reflect late maturity in Finland, or the results might be affected by companies' adoption of I-EDI at that time; either way, there is no data available that would explain the trend.

Unfortunately there is no statistical data available to assess the adoption rate in the 1980s and 1990s but the literature suggests that the majority of adoption took place at that time. An interviewee with a long history from the financial sector explained:

> *It really began in 1985 when we began building systems to use data in real time. At that time I created a program calculating the Finnish Mark index which utilized the latest innovation in the industry, data feeds...*

He also explained in length how important that data was to the business — it allowed the bank to operate more efficiently due to the reduced arbitrages. The reason this is interesting is because it is already a very primitive example of an API providing a real time data feed to a company. However, at the time the system was immensely complex and required significant set up and installation to provide the single data feed.

In general, the adoption EDI systems were largely associated with the adoption of computers in the companies — the reasons and benefits of adopting them were largely related to being able to access crucial business information efficiently (e.g. currency indices or credit risk) and the interviewees largely highlighted them being used for inbound information and data rather than sharing it externally. Hence, the value of EDIs and the adoption reasons were largely associated with streamlined operations and reduction of manual labor, and primarily for internal operations and reasons. No interviewee mentioned sharing information or data with external parties using EDIs unprompted; such practices were likely conducted at the explicit request of some external partner.

Although it's a surprising empirical result when reflecting to the literature, which suggests information sharing was often a dyadic relationship, the difference can be explained by the sample of experts interviewed. However, it still implicates that EDIs were adapted primarily for the purpose of improved operational efficiency, regardless of the firm-specific use cases — no strategic reasoning can be found in the literature nor from the empirical data used in this study.

4.1.2 Adoption of internet

The literature suggests that internet drove the adoption of data sharing systems to much smaller companies due to the lowered adoption costs (Zhu et al., 2006; Huang et al., 2008). In figure 5 we can see the use and penetration rates of web-based systems in comparison to their EDI counterparts.

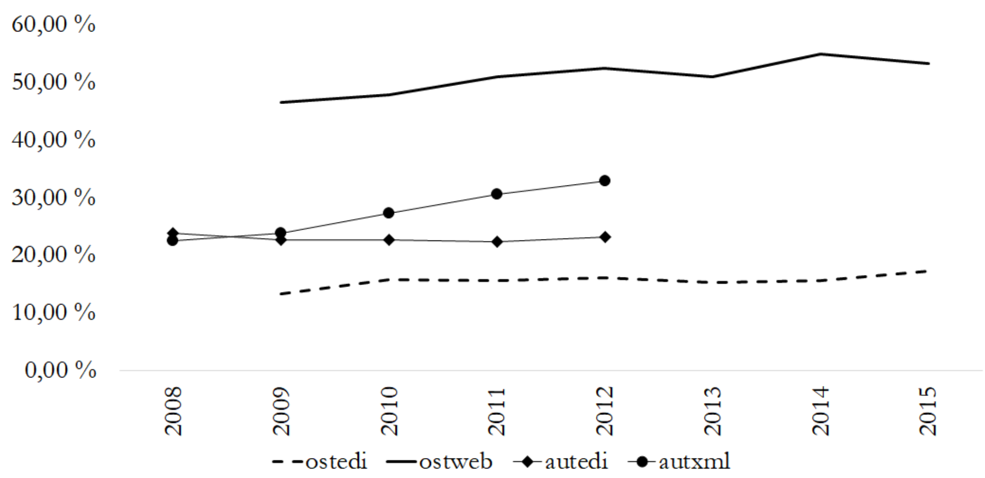

Figure 5: EDI to web transition

In 2008, the automatic data exchange was conducted using EDI and XML in relatively same proportions (23.9% vs. 22.6% respectively) with EDI being used in a slightly larger share of companies. However, after four years in 2012, XML was used in 33.0% of the companies whereas EDI was used in 23.3% — XML use rate had grown 9.9% annually while EDI use rate had decreased slightly. This data strongly supports the literature and the hypothesis that open standards were adopted very quickly in comparison to EDIs, possibly driven by the low costs of adoption and the subsequent low-friction access to these technologies.

The other comparison in the same figure 5 is between the used standards in companies' buying orders in 2009–2015. We can clearly see that web-based technologies have penetrated over half of all companies in this regard (53.3%) while EDI usage remains quite small in 2015 (17.3%). Interestingly, the EDI usage has grown in the given timeframe from 13.4% in 2009, indicating the importance of maintaining the EDI systems in certain companies and supply networks.

Although the interviewees had difficulties distinguishing EDI and I-EDI systems from each other, and therefore explicitly comment on this change, all interviewees recognized internet as something that the companies adopted rather quickly. For example, one of the interviewees saw their organization of tens of thousands of people change from no internet access to every function of the organization connected to internet in 6 years.

> *...it's difficult to imagine that they would have had a real integrated weather service. I think that it must have been more like, an analyst looks it up somewhere, and gives inputs to the algorithm. I think they utilized data quite widely considering the time [2006]. (Interviewee worked in the same company, but another team)*

This quote highlights the technological limitation to the use of external data feeds to complement the analytics. The focal company was, according to the interviewee, advanced in its industry regarding their data exploitation and analytics practices. They had recognized the need to use a wide variety data to generate better insights, but there were no open APIs (or equivalent) available to efficiently have access to the desired data.

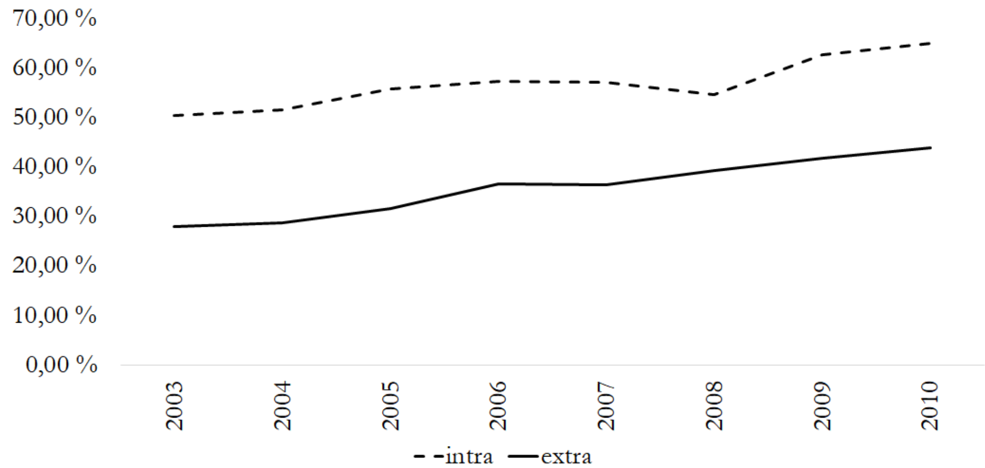

Figure 6: Comparison between intranet and extranet adoption

Considering companies' adoption of extranet, we can see from figure 6 that between 2003 and 2010 intranets were more common in companies, suggesting that the adoption of extranets could have been considered primarily as an extension of intranet. However, the significantly larger compounded annual growth rate of extranet (6.6%) to the one of intranet over the same time period (3.7%) does suggest that extranet was considered as more important to many companies than intranet. This finding is in line with the hypothesis that there might have been network effects at play which drove the adoption of web based data sharing systems, such as extranets, more quickly than would otherwise have happened.

The impact of the public web disrupted the research and development in some industries and fields quite drastically. An interviewee explained how web-based repositories were implemented to replace physical distribution of test results and allowed a much wider range of parties to participate in the development of technology standards. He also pointed out how the inclusion of a greater number of parties resulted in much quicker development and standardization processes.

4.1.3 Modern era and APIs

Of the 1303 firms surveyed by The Research Institute of the Finnish Economy in 2016, only 83 companies (6.4%) stated that they have an open API. It highlights the early stage of companies' adoption of such modern data sharing technologies. An expert also said that although insurance companies buy a lot of data nowadays, they are often delivered in large batches and many of their data suppliers do not necessarily offer APIs for the data feed, which supports the low penetration 6.4% suggests. No interviewees contradicted the consensus that the relevance and availability of APIs will be greatly expanded in the coming years.

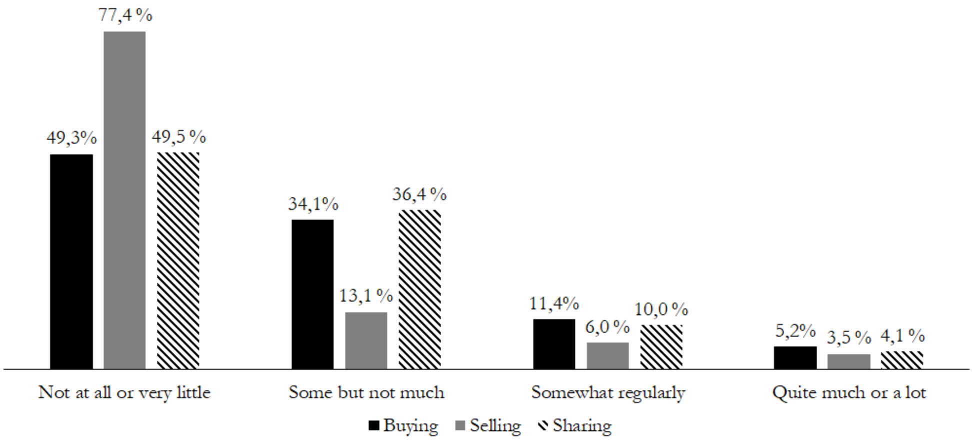

Figure 7: Surveyed companies' big data practices

In figure 7[1] we can clearly see how the surveyed companies recognize the value of big data, as over 50% of respondents buy more than a little big data. However, the difference to selling big data is significant suggesting that companies have not yet integrated big data to their sales. Additionally, sharing of big data is also less common than buying it — the factors together suggest that companies are primarily leveraging big data internally and realizing only operational efficiencies and some product improvements.

> *If we are talking about banking sector . . . PSD2 will be a big change because it forcibly opens our APIs outwards and subsequently opens up opportunities for our business.*

Interestingly two of the interviewees highlighted the role of regulations and legislation on the emergence of open APIs. The aforementioned PSD2 directive, according to one expert, could be a prelude for a widespread phenomenon requiring companies to provide APIs to their data.

When discussing the way in which a social media analytics company delivers their insights to companies, the interviewees said:

> *. . . they offer an API, and they do not send any of the stuff via an e-mail any more. They gladly offer and API to which the*

[1]Subset (n=301) of the respondents who answered that they use big data in their business

> *company integrates to, after which it gets the already analyzed end result into their own systems via integration, which can then be attached to some reporting for example.*

This quotation has two interesting dimensions and insights; firstly, it goes to show how new business models have been created around APIs as the social media analytics company actively monitors social media feeds via APIs, refines the data into actionable insights, and provides them to a larger industrial company via an API that can be integrated to a dashboard for e.g. marketing department. Secondly, this example shows how companies are connected to an increasing number of partners that provide value to their company in various forms — the company did not have a business relationship with social media analytics company a decade ago, and now it is actively transacting with them.

> *Especially with external partners, we also started to demand them that they need to have a proper API, because we did not do it any more so that an Excel is sent via e-mail, where the data is, but there has to be some proper API.*

As highlighted by the interviewed IT expert, the forerunners of adoption are already starting to demand their partners to provide proper APIs as a method of data sharing or transfer, a similar practice that was observed in the transition to I-EDIs in 2000s (Bednarz, 2004).

4.1.4 Main findings

The quantitative analysis of Finnish companies has shown how the technical capabilities and the systematic integration of data into organizations has accelerated, and the interviews suggest that it is due to external pressure and the increasingly attractive benefits that can be realized. In just two years (2014–2016), the share of Finnish companies systematically using data to improve their business grew over 20%. Such rapid adoption cannot be observed from other time periods available in the data of Statistics Finland.

While the adoption of technologies and practices has grown in the past years, a relatively small portion of companies are taking advantage of the possibilities. Only 6.4% of surveyed companies have an API and only 23.3% of surveyed companies[2] use big data in their business while all interviewees agree that using big data enables benefits of some sort in virtually any industry. Such a discrepancy suggests that there are locked benefits for companies

[2]The sample may be biased to show too favorable numbers

that require relatively small investments into IT assets, as the costs of adoption have shrunk significantly.

As literature suggests (e.g. Liozu and Ulaga 2018; Moilanen et al. 2019) APIs are relevant to any firm's business to participate in data economies, and companies are already starting to demand proper APIs from their partners. All expert interviewees highlighted the disruptive potential of PSD2 to banking business — they offer opportunities to start building ecosystems around financial services and creating new businesses to exploit the new opportunities. Ultimately, as one interviewee stated, it likely leads to superior customer experiences that can also result in improved performance. There are already companies, such as the mentioned social media analytics company, whose products or services rely on APIs and their own product might also be delivered via an API.

The interviews also validated the hypothesised evolution from EDI to API, and when prompted they agreed on the connection. The transition from EDI to I-EDI is a clear evolution, and can be seen from both the quantitative data (e.g. figure 5) and from the way interviewees differentiated the time before the internet to the time after. The evolution from I-EDI to API is not necessarily as rigid, but the way they are treated nowadays validates them as being in the same linear progression — from extranets and SAP to open APIs easily accessible via any software.

4.2 Role of data on firm's performance

Unfortunately, no quantitative data regarding the benefits of data sharing was available for this study. Liozu and Ulaga (2008, p. 151) provide a summary of some of the quantified benefits in the literature and individual studies are referred to in chapter 2 (e.g. Mukhopadhyay et al. 1995). Hence, this section is largely driven by the expert interviews.

However, data from Statistics Finland innovation survey helps understand how data is currently used in companies and how it has changed between the two observation years in 2014 and 2016. The share of companies using data in a given dimension can be seen in figure 8.[3]

[3]For variable names used, refer to table 8 on page 39.

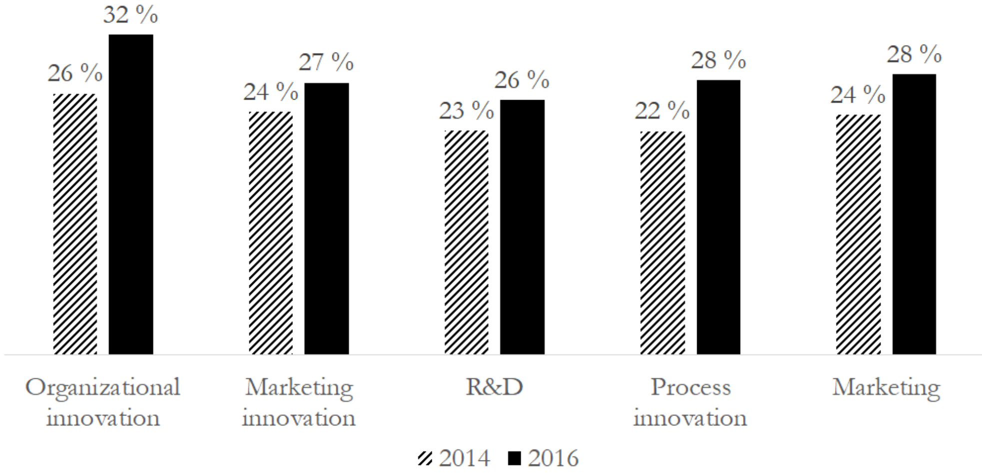

Figure 8: Cases in which Finnish companies utilize data

The most common use cases for data use are in organizational innovation (32.1%), marketing (28.1%), and process innovation (27.6%), while the use of data in research and development lags behind (25.6%). These numbers suggest that data is used primarily in optimizing certain aspects of the business more than it is used in order to seek new opportunities — an insight that came up on numerous occasions in the interviews where the primary motivation for use of data has been to optimize various aspects of the business and cut on inefficiencies. When looking at the compounded annual growth rates of these percentages, use of data in process innovation has grown the fastest at 10.9% annually while the same growth in research and development has only been 6.6% annually. Such development suggests that during the observed time frame 2014–2016 companies are indeed focusing on optimization and reducing inefficiencies. However, interviewees also confirmed that in their experience companies have very recently started to recognize the potential and the adoption is underway — unfortunately, as of July 2019, Statistics Finland has not released the results of 2018 and therefore a large scale shift in companies' attitude cannot be quantitatively validated or confirmed.

4.2.1 Readiness to share data externally

In order to be able to capture the benefits of data in the external dimension, companies have to share the data externally as well.

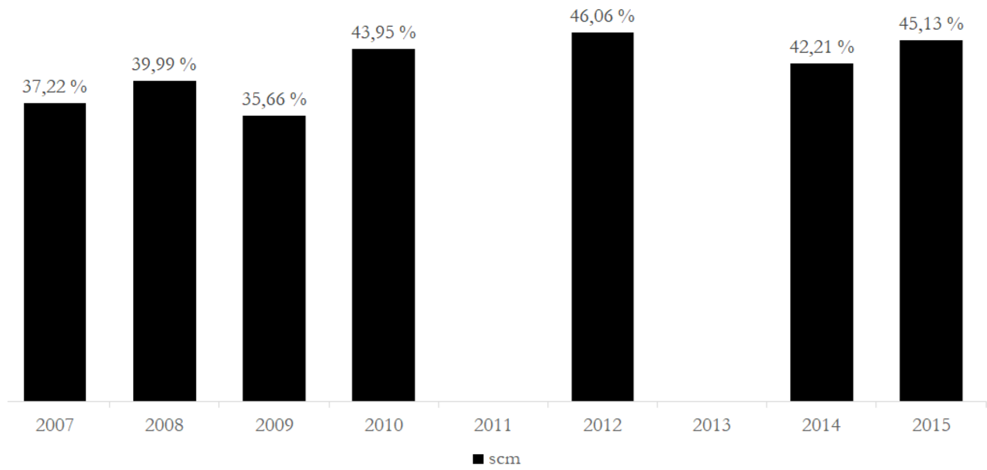

Figure 9: Share of companies sharing information with their supply chain partners

As discussed in chapter 2, information sharing has been common practice in companies for decades already. Figure 9 shows the share of Finnish companies that share data in their supply chain — we can clearly see that a significant share of companies (45.1% in 2015) already share this kind of data, which indicates that companies are accustomed to data sharing in some contexts at least. From the same figure it is clear that companies are more accustomed to buying big data than selling it — 77.4% of companies state that they do not sell big data at all, or do it very little while the respective number for buying big data is 49.3%. Although the numbers differ, likely due to sample size and selection bias[4], data from Statistic Finland supports the same conclusion; in 2016, 23.7% of all Finnish companies stated that they buy data externally, while only 16.9% sell data externally. What is important to note is that the same shares were 19.9% and 14.1% respectively in 2014 indicating an annualized growth of over 9% in both dimensions, suggesting a relatively rapid adoption of data business in companies.

Of the 303 Finnish companies that responded that they use big data in their business, only 21 companies (10.2%) responded that they do not sell, buy, or share big data at all with external parties, but rather only use data that they gather themselves. This suggests that almost 90% of companies using big data to some extent actively participate in data trading or sharing which strongly supports the assumption that the emergence of big data and

[4]The respondents to the survey by The Research Institute of Finnish Economy were companies that were already involved in data business in some form. Additionally, the responses were voluntary as opposed to the mandatory nature of the Statistics Finland's surveys.

widespread real time data sharing has created new business opportunities. One also interviewee predicted a significant business opportunity for companies that will be able to play the role of a data aggregator as the data trading and sharing networks become increasingly fragmented.

Two interviewees highlighted the importance of ”trusted partners” i.e. emphasizing the aspect that data sharing remains rather closed and constrained to inner circles[5]. However, the following quote gives an alternative viewpoint in the context of R&D:

> *Quite far it [data sharing] was done for cost optimization purposes. But there was also that we could create a kind of a community who collaboratively created the algorithms and then we could create the ”best of the best” in standardization.*

The interviewee explained that such thinking was already very common in the late 80s when they were developing standards regarding telecommunications and radio frequencies, and that the scope of the collaboration was greatly intensified when 3G technologies were being developed in the early 2000s.

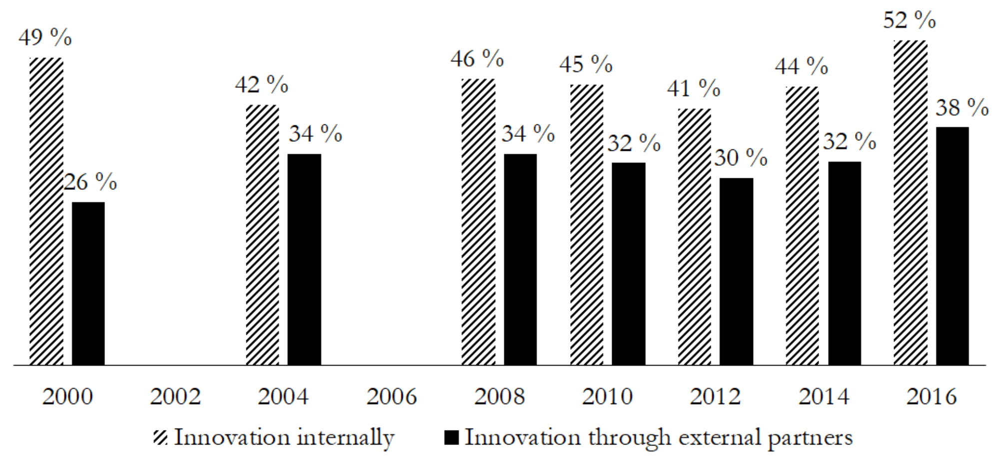

Figure 10: Share of companies that have done product of service innovation internally and externally

Figure 10 validates quantitatively how the share of companies working with external partners in their innovation practices is quite high as in 2016 approximately 38% of companies had done product or service innovation

[5]For academic discussion regarding data openness and ownership, see Rajala, Hakanen, Mattila, Seppälä, and Westerlund (2018)

through external parties and the percentage has grown 2.4% annually since 2000, while the purely internal product or service innovations have grown only at 0.3% CAGR indicating a trend towards joint innovation which involves data sharing of some kind.

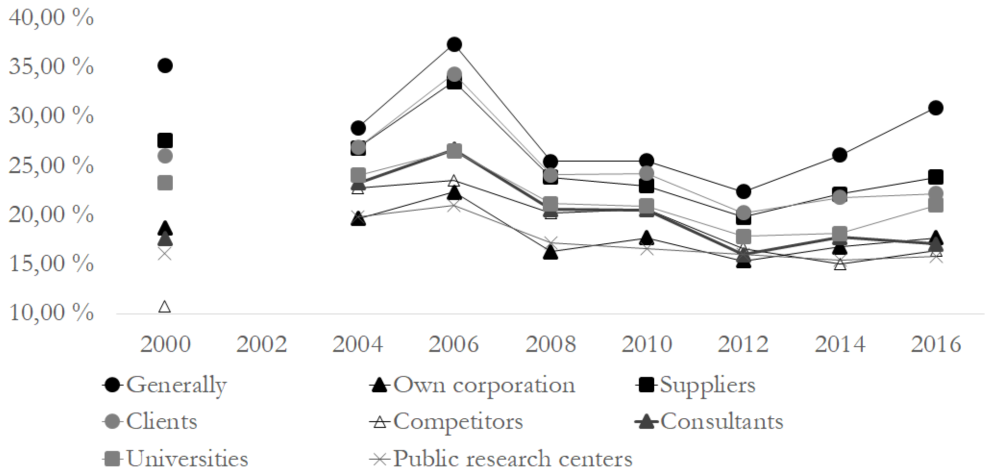

Figure 11: Share of companies doing cooperative innovation with various entities

Figure 11, on the other hand, suggests a different story although the results are from the precisely same survey by Statistics Finland. In this figure, the shares of companies that are doing cooperative innovation with a specific entity are listed annually. According to the responses to this data, 35.2% of companies were doing cooperative innovation in 2000 while the comparative share is only 30.9% in 2016. The only growth in the same time frame is seen in cooperation with competitors, which grew from from 10.8% to 16.4%.

Cooperative innovation itself was not assessed in the interviews but it came up but joint initiatives came up in two; either in Finnish banks creating their shared credit card issuer (rather than using bank-specific cards) or in insurance companies sharing claim histories together in order to assess customer riskiness.

4.2.2 Operational efficiencies

All of the interviewees were unanimous in emphasizing operational efficiencies being the primary driver for data business—trading or sharing—as means to cut inefficiencies in operations.

One of the interviewees worked at Finnair, an airline company, in 2006–2008 and perceived the analytics and data being more advanced than the industry average. While all airlines have incorporated capacity planning and forecasting into their processes, Finnair also used other data sources e.g. calendar events and weather data. However, the following quote provides a tangible example of how having access to the relevant data transformed the entire way of running business operations at Selecta, an automatic coffee machine company in 2009–2010.

> *There was a lot of effort put into integrating the company internally, and the big driver was that the coffee machines ... were entirely offline, so the firm does not know how much there is coffee and if it works. Then some human must run around in a predetermined route to fill the machine, then the one in the next building — sometimes there was no need to fill the machine. ... then we started to develop these IoT-based solutions that worked over the internet. .... Once we could do this, it entirely transformed the business.*

In essence, this is a clear example of realized internal operational efficiencies. Having access to real time data of coffee machine water level, possible malfunctions, and the amount of coffee beans left drastically reduced the amount of labor needed for their upkeep as the company started to do dynamic route planning for the employees. Additionally, the link to technology is clear as the interviewee explained that they retrofitted their coffee machines with 4G modems to enable the real time data transfers — indicating that such practice would have been infeasible without such technologies. However, the data was held tightly within the company with no data trading activities implemented into any direction with the only exception being the outsourced maintenance staff that could monitor the machines' status in real time. At that time, the process was being transitioned to predictive maintenance to service the machines before customers would have to call the company and file a complaint or service request.

4.2.3 Strategic opportunities

According to the experience of the interviewed IT expert, the use of data and analytics for dynamic pricing and targeted sales started around 2013. Similarly, the experts from financial sector highlighted the importance of data in being able to sell more products to customers due to being able to reach them at the right time and with the right offer. Reflecting to the categorization of the benefits, these can be considered as strategic opportunities as they enable new products and services and revenues that may not have been available previously.

> *It was seen that if we would be able to provide this kind of platform, on which they [mobile apps] could be built, that would work across company borders, then it would lead to quite a lot better customer experience. .... Therefore, the driver [for adoption] is purely, by customer experience, and genuinely a good customer experience, new business.*

The quote above is a description of why the insurance company, in which the interviewee was working at the time, was interested in exploring options in opening up APIs. In his experience, banks were still quite conservative in exploring strategic opportunities enabled by APIs—as mentioned by two other interviewees as well—while insurance companies were already actively looking for such opportunities. When asked further, he estimated that no significant benefits have been realized despite numerous pilots but explained that it could be due to the amount of data a company has to accumulate before it's sufficient to build an ecosystem on.

In addition to speculative opportunities, data has also provided companies with tangible strategic benefits, such as a bank quickly reacting to shifting interest markets with a long-term locked interest mortgage, which became a major success for the bank. It was successful because the bank executives had real-time data of the market and could run the necessary analytics and design an impactful marketing campaign all within a week, being the first one to the market.

Analytics and big data are increasingly being incorporated into businesses and above are some examples of how they are being used in the financial sector in Finland. Globally, the opportunities are much wider:

> *Like in the USA they sell it [data] just like "hey, here are five million people for you that we think are, for your industry and company, useful target market because they behave like this". Such things I have not heard of in Finland.*

This emphasizes the possibility that the data sharing and trading practises significantly vary depending on regulations and culture in any given geography. However, it also provides an opportunity to learn from companies that have been successful in countries where such actions have been possible.

4.2.4 Main findings

Clearly, the main motivator for using and sharing data in companies has been to achieve operational efficiencies or to drive costs down. More recently, opportunities have been identified in generating more revenue from customers by providing better products and services with the help of data.

Companies are accustomed to sharing data, but primarily with selected partners and primarily specific predetermined data with the intention of gaining direct benefits from the activity. Quantitative analysis supports the literature by showing that approximately 45% of Finnish companies already share data within their supply chains to realize the resulting operational efficiencies, with the number increasing as only larger enterprises are considered. However, only 14% of surveyed companies share big data externally somewhat regularly or regularly, suggesting that the external strategic value of data is still largely unknown to companies. Interviews also support this notion; companies have identified the shift in the market as various platform businesses succeed, and executives are considering the implications to their companies. The actions remain quite small, and the companies remain in exploratory mindset. PSD2 affects all banks which has fuelled the discussions regarding these discussions, and one of the interviewees illustrated how his current employer is looking to create a financial services platform around the open APIs enforced by PSD2. They are not yet sure on how to monetize it, but it serves as an example of companies taking initiative to explore the opportunities.

The following quote from the publication event of Digibarometer[6] 2019 highlights how a large Finnish forestry company representing advanced use of data exploits it:

> *More than 80% of all digitalization initiatives are focused on operational efficiency [and 20% on developing new products and services] because it is much easier to work with what you have already. .... The value of the project has to be significant for us to work on. (Pasi Kulmala, Stora Enso)*

The quote is well in line with the already discussed notion that companies are primarily focusing on the operational efficiencies enabled by digitalization and data. Kulmala explained that it is natural because the impact of cost improving innovations are quantifiable while developing new products

[6]Digibarometer is an annual research that ranks countries based on their utilization of digitalization, and it is conducted by The Research Institute of Finnish Economy. Recording of the event can be found at `https://tv.streamfabriken.com/digibarometri-2019-julkaisuseminaari` (1:01:10) (in Finnish). Accessed 18.06.2019

Dimension	Sub-dimension	Example
Strategic opportunities	Internal	Being able to quickly develop a new financial product to meet the new needs as interest markets quickly shifted. A desirable product ready for sale within a week of the first indicator
	External	Hybrid products including insurance products with certain mortgage customers
Operational efficiencies	Internal	Adoption of web banking resulted in more efficient branch operations and opportunity to close some branches down
	External	Utilizing external parties' analytics (e.g. social media) to focus efforts

Table 10: Examples of data benefits in financial sector provided by an ex-CEO of a bank

and services is unpredictable — initiatives with too small business potential are not pursued. The access of third parties to develop these products and services is also tightly controlled which limits the opportunities of smaller companies, who would be willing to take the risk, from developing truly innovative products and services that could—hypothetically—generate significant revenues to Stora Enso when aggregated.

The framework for assessing the benefits of data also proved itself a valid and intuitive lens for its purpose. One of the interviewees, an ex-CEO of a large bank was eager to provide some examples of the benefits of data directly referring to the framework on p. 30. The results can be found in table 10. Additionally, the benefits discussed in the expert interviews as well as presented in the Digibarometer 2019 event could all be placed into the framework.

Chapter 5

Discussion

Product modularity has been recognized as being beneficial to the products'—and subsequently the respective companies'—performance (Antonio et al., 2007) and the study has shown that modern technologies allow product modularity to exist in the software realm as well. By providing sufficient boundary resources—such as APIs, SDKs, and access to company data—to external parties enables companies to tap into external strategic opportunities of data economy.

The existing literature does not assess APIs to be an evolution from EDIs, but instead address the APIs as a separate phenomenon linked to platform business (Jacobson et al., 2011; Moilanen et al., 2019) aside from one unpublished paper (Benzell et al., 2017). This study has successfully contributed to the literature by laying out the foundations of API by linking it to previous technologies and verified the logic with experts. Due to this differentiation in the literature, the discussion revolving around the benefits of data was also separate — the real time exchange of information resulting in operational efficiency and new data-driven business models and complementary innovations on the other hand. By assessing these separate points of view and combining it with the technological advancements with an abductive approach, a unified view is proposed.

By triangulating theoretical background with quantitative and qualitative research, this study has succeeded in linking the use of data sharing technologies to tangible benefits for companies. While there are large factors moderating the effect—the industry, competition, and executive readiness to name a few—the potential benefits of adopting sophisticated data sharing technologies and practices should be noted by all companies.

While the notion of boundary resources[1] is often linked to platform econ-

[1]See page 7

omy and considers APIs to be a form of boundary resource, we can consider all data sharing technologies as being a form of boundary resource through which external parties are able to utilize to various degrees in their own business. This study contributes to the boundary resources discussion by proposing tangible benefits available through treating information systems and data as boundary resources to external parties.

This study has shown how expanding openness, and involvement of external parties through data sharing, has enabled progressively more significant benefits through the last few decades. By electronically sharing timetable and inventory information with key partners, companies were able to reduce the need of labor internally already in the 1970s (Zhu et al., 2006), and decades later providing open access to software platforms through various boundary resources (Parker et al., 2016b) companies have been able to monetize on complementary innovations that can already generate a significant portion of the companies' revenues. As technology progresses and information will be further commoditized, even more significant opportunities could emerge.

Information and data sharing is not a new phenomenon to companies as it has been practiced to various extents across all industries. Hence, the decision to expand these initiatives is merely a matter of changing the business mentality; the executives should start asking themselves how they could take advantage of external parties through APIs, or what kind of data do they possess that could be valuable if exploited correctly. This allows more companies to adopt platform-like business models that revolve around data or allows businesses to realize similar efficiencies to their best performing peers.

In chapter 3, three hypotheses were introduced from the literature. The first hypothesis can be largely confirmed; statistical data shows how XML-based technologies exceeded the EDI technologies in the early 2000s. However, the penetration and exploitation of APIs cannot be exactly determined due to the lack of data, but expert interviews suggest that they are at least discussed across companies and industries in some level. The second hypothesis can also be confirmed; specifically, the internet and open APIs have allowed companies to work with more parties and develop their offerings in ways that were not possible prior to having access to the data they have nowadays. IoT solutions and integration of various external data feeds are nowadays used for the development of better products and services. The third hypothesis remains valid, as it could not be contradicted with empirical evidence, but it should be researched more; interviewed experts' experiences could be categorized within the framework but that alone is not sufficient evidence of the robustness of the framework. Further research is suggested in order to tweak and validate the categorization of data benefits. In chapter 1, three questions were introduced in order to understand what kind of performance-related benefits does data enable in companies with modern technologies, and below are the summarized answers based on the analysis.

How has the evolution of information technology affected the sharing and use of data and information?

Before the adoption of internet, interorganizational information technology systems were primarily used by large corporations who had the resources and the capabilities to adopt EDIs to effectively share operational information within their supply chains (e.g. Takac 1992). The systems had limited applicability due to compatibility issues with other systems and were used to transmit specific operational data, such as orders or inventory levels. These practices allowed companies to improve their operational efficiency by e.g. reducing buffer inventories and the amount of manual labor involved in handling and transmitting the information.

The emergence of internet and web-based information systems changed the dynamics of data use and sharing in companies; an increasing number of companies, including small and medium-sized businesses, were able to link their systems to the internet with little cost (e.g. Huang et al. 2008). In Finland, the statistics show that by 2009 the use of XML for automatic data exchange surpassed EDI based technologies and the adoption grew quickly afterwards. Protocols were standardized and companies were able to form links with partners who had not previously adopted the EDIs. Larger companies developed extranets that enabled external parties to access the relevant business data autonomously with proper credentials and the need for labor-intensive integration work was cut.

APIs enable a virtually limitless number of parties and connections into systems with very little integration work (Jacobson et al., 2011; Moilanen et al., 2019). They have created a new economy in which companies do business by providing APIs for other companies that e.g. need the data in their analytics. The full extent of the possibilities enabled by APIs remains unclear but interviews suggest that APIs will be a key consideration for businesses across industries as data and analytics are increasingly in focus of businesses.

What kind of benefits does the use and sharing of data entail in companies?

The benefits enabled by using and sharing data in companies can be categorized into operational efficiencies, focusing on cost reductions, and strategic opportunities that enable companies to generate more revenue. No similar framework was found in the existing literature, although the various benefits can be categorized into these; as for tangible examples, Tao et al. (2018) outline how selling analytics with manufacturing machinery has enabled new business opportunities, and Davenport and Kudyba (2016) explain how sufficient data coupled with predictive analytics can reduce the maintenance and upkeep costs of various assets.

Sharing data with external parties enables external benefits as well—creating the internal and external sub-dimensions into the framework—as both the operational efficiencies and strategic opportunities can be developed outside the focal company. As for external operational efficiencies, ample literature exists on the benefits of information sharing in supply chains (e.g. Stevens 1989; Gavirneni et al. 1999; Cachon and Fisher 2000; Lee et al. 2000; Lee and Whang 2000; Yu et al. 2001; Lau et al. 2003; Vickery et al. 2003; Barratt and Oke 2007; Prajogo and Olhager 2012; Lotfi et al. 2013) and the expert interviews suggest that the ability of external companies to provide analytics as a service, improving the focal company's performance, can also be categorized as an external operational efficiency. The external strategic opportunities remain largely theoretical—aside from a few examples—but in principle they stem from the opportunities enabled by product modularity through software, and from complementary innovations. Although not discussed from these lenses, literature exists on collaborating with externals (Tapscott and Williams, 2006), on complementary innovations in traditional industries (Peppard et al., 2011) and more notably in platform business (Parker et al., 2016a,b; Zhu and Liu, 2018). According to the interviews, exploiting these external strategic opportunities is an increasingly attractive proposition for companies, but very few have the necessary capabilities and knowledge to integrate them into companies' strategies yet.

How could the benefits of data-related activities be quantified?

By assessing the benefits related to information technologies, data sharing, and product modularity, a unified framework for data-related benefits was developed. The framework remains quite abstract, but its applicability was verified with multiple experts and all benefits from both empirical data and existing literature can be mapped into this framework. Hence, the total monetary benefits available from data-related activities could hypothetically be calculated from:

$$\sum o_i + \sum o_e + \sum s_i + \sum s_e \tag{5.1}$$

Dimension	Example literature
$\sum s_i$	• Sales from new data / analytics products (Woerner and Wixom, 2015; Liozu and Ulaga, 2018) • Increase of sales in better customer targeting (Constantiou and Kallinikos, 2015)
$\sum s_e$	• Revenue generated by external parties on the platform (Tiwana, 2008; Parker and Van Alstyne, 2018) • Preliminary results show that opening APIs improves market value (Benzell et al., 2017)
$\sum o_i$	• Reduced labor costs (Mukhopadhyay et al., 1995; Iskandar et al., 2001) • Streamlining processes (McAfee and Brynjolfsson, 2012; Davenport and Kudyba, 2016)
$\sum o_e$	• Supply chain efficiencies (Lee et al., 2000; Cachon and Fisher, 2000; Lau et al., 2003) • Utilizing external parties' expertise in analytics (Davenport and Kudyba, 2016)

Table 11: Possible levers for quantifying the impact of data

In equation 5.1, o_i denotes the internal operational efficiencies, o_e external operational efficiencies, while s_i and s_e denote the internal and external strategic opportunities. Hypothetically, the sums of o_i and o_e would result in cost savings while the sums of s_i and s_e consist primarily of revenue increasing components. Further studies should be conducted in order to break the components down for practical use — Liozu and Ulaga (2018, p. 151) provide some of the operational efficiencies and their drivers, but no literature exists that has provided ways to quantify the revenue increases of data economy. This study has been able to qualitatively propose the mechanisms and core principles why the cost reductions and revenue increases exist and firm-level case studies are suggested in order to start quantifying the model.

In table 11, the elements for quantifying the whole impact of data on firm performance has been tied to some examples of the reviewed literature to illustrate the kind of components that should be used in the different categories of benefits. The listing is not exhaustive and warrants future research to generate a comprehensive list of benefits and means to quantify them individually. To quantify the operational efficiencies enabled by data—although complex and difficult to estimate (Brynjolfsson et al., 2002)—vast literature exists since 1990s when EDI systems were being implemented in first companies, and supply chain efficiencies were being identified from information and data sharing practices.

As for strategic opportunities, literature is much more recent and scarce. The value of analytics is being discussed in making better decisions (e.g. McAfee and Brynjolfsson 2012; Chen et al. 2012; Davenport and Kudyba 2016) but they are addressed separate from each other rather and no literature combines the empirically identified levers, such as better sales efficiency and customer targeting, with the benefits of analytics in strategic decision making. In the external dimension, this study has identified the opportunity to leverage platform and software modularity—and the related complementaries—by laying out the theoretical basis for the argument, and verifying the existence of these opportunities with expert interviews. Unfortunately, very little research exists in this dimension, and hence most of the levers remain unknown — yet, Benzell et al. (2017) in their currently unpublished study found that opening APIs positively affects companies market value. In future research, literature on multi-sided markets could be examined (e.g. Boudreau 2010; Smith et al. 2016; Parker et al. 2016b; Parker and Van Alstyne 2018) to explore the external strategic opportunities since their business models are largely built around these ecosystems and revenues generated by third parties.

Implications

Although businesses across industries are increasingly connected, many companies have failed to grasp the opportunity and rethink their business models as only approximately 23% of surveyed Finnish companies use big data in their business. Data and analytics are primarily used to achieve superior operational efficiencies i.e. making current business more productive, while only some companies have successfully adjusted their business models to harness the opportunities and create new business altogether. Ample literature exists on harnessing the opportunities (e.g. Jacobson et al. 2011; Constantiou and Kallinikos 2015; Davenport and Kudyba 2016; Liozu and Ulaga 2018; Moilanen et al. 2019) and the results of this study suggest that companies largely already possess the technological requirements and it is primarily a matter of changing the managers' perspective and incorporating data economy to the companies' strategy and start discussing the opportunities of industry platforms.

For descriptive purposes, the results can be reflected on the platform classifications by Gawer (2014); companies have the technological capabilities and infrastructure at hand to utilize these technological platforms but the implementations remain in internal or supply-chain platforms at large. This enables the companies to exploit the innovative capabilities within the firm, partially within their supply-chain network, but they fail to tap into the potentially unlimited pool of external capabilities (Gawer, 2014).

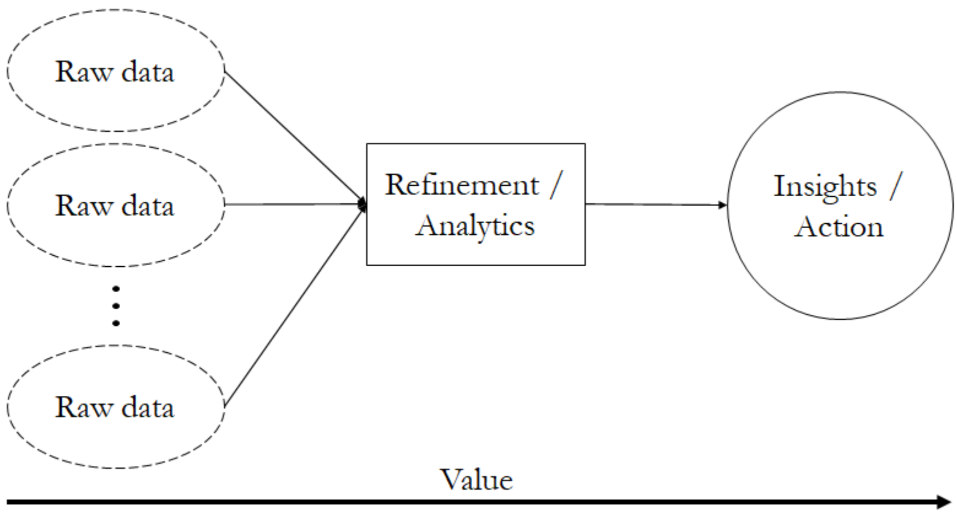

Figure 12: The value of data comes from the analytics and insights derived from it

It is important to note that raw data is not valuable in itself, as noted in both literature (e.g. Liozu and Ulaga 2018) and in expert interviews. Data—including big data—is valuable when it results in some insights or action in e.g. strategy making (as discussed by e.g. Constantiou and Kallinikos 2015; Woerner and Wixom 2015) or optimizations (e.g. Lee and Lim 2003; Najjar and Kettinger 2013). The empirical data of this study suggests that the data-driven benefits are not only theoretical — an interviewee gave an example in which data-driven decision making lead to the introduction of a new financial product that had not been available on the market before. Raw data alone is valuable to those parties that are able to either refine it or derive some useful insights out of it by e.g. combining it with other data sources. This is largely enabled by modern technologies; data has evolved from being an instantaneous electronic message improving operational efficiency to being a resource in its own right with its own strategic value and implications. Hence, companies considering the implications of data economy and data monetization must look beyond the raw data in order to understand where the value can be created. Similarly, purely the opening of an API is not enough to realize tangible benefits; it is important to assess what kind of data is available through the APIs and to provide sufficient documentation and support to the developers in other companies willing to use it.

5.1 Limitations of this study

In this study, the role of data on firm performance has been studied by combining the technological and business perspectives, but the phenomenon also encompasses perspectives from knowledge management[2], and other social fields of study. Additionally, much of the literature addressing the impact of APIs, big data, and analytics is still quite early due to the early nature of the subject. Hence, the generalization of the results of this study should be avoided without further research.

The quantitative methods used to address the technology adoption rates and readiness to share data with external parties had to be conducted with several limitations. There is no data available that would address the questions specifically, and several assumptions and interpretations of the answers had to be made. Fortunately, the conducted interviews support the insights derived from the quantitative analysis but to improve the rigidity of these results more detailed and comprehensive quantitative analysis should be conducted. As for the quantitative impact of data on firm's performance, a study

[2]For more discussion on information sharing, see e.g. Alavi and Leidner 2001.

on USA-based companies—which have been able to buy e.g. customer purchase behavior data from credit companies—could be conducted because the regulatory environment has been much more relaxed in comparison to Europe or Finland specifically

Additionally, while the triangulation methodology is good for theory development due to multiple perspectives on the same issue, it cannot guarantee the validity of any given theory. For example, many of the experts interviewed for this study had background from the financial sector due to the ongoing conversations driven by PSD2 and its implications. Hence, the developed theory must be tested in a future study in order to validate the proposed theoretical link between data sharing and realized benefits through more comprehensive review of different industries and experts.

5.2 Suggestions for future research

Since the end result of this study is largely theoretical, the results of this study must be tested in practice. Particularly the quantification of data related benefits should be researched in detail, possibly via several case studies with cooperating companies. Additionally, the possible implementation challenges and considerations should be explored in more detail before recommending all companies to take action and seek external strategic opportunities with their data — interviews suggest that management capability and motivation can significantly moderate the success of such initiatives, but such considerations were left outside the scope of this study.

Although the categorization of data related benefits was developed from theoretical basis and validated with experts, it should be tested with larger sample of empirical data. The interviewed experts in this study are Finnish and largely represent the financial sector, which might distort the results to some extent.

Considering the evolution of modularity from products to services, and further to platforms, no literature exists to consider the modularity of data or the required interfaces and dynamics. Some of the interviews suggest that data is becoming an increasingly valuable asset on its own, and hence it could be hypothesised that it could be considered from the lens of modular design as well.

Chapter 6

Conclusions

The advancements in digitalization and information technologies has allowed most progressive companies to monetize on new assets altogether by providing software and data platforms on which complementary innovations can be developed. The opportunities are closely tied with the notion of product modularity — a few decades ago products were developed to be compatible with other hardware, but now any party can develop new software products and distribute them with little cost.

Statistical data in Finland suggests that adoption of data sharing technologies has accelerated with the evolution of technology. This has allowed an increasing amount of companies and individuals to take part in the emerging ecosystems due to accelerating network effects. The role of these technologies has been important in enabling companies to realize the benefits of having access to an expanding pool of data — both by having all data internally available and by being able to utilize a large number of data sources and partners simultaneously to build increasingly complex analytics and algorithms. Interviews also highlighted how the reason for the technology adoptions has also changed with time; the EDIs were implemented to make an individual process much more efficient, internet-based systems were adopted for more holistic connectivity and the industry driven peer pressure, and APIs are seen as necessary parts of future business models. Hence, the proposed framework linking technological advancements to benefits remains uncontested.

The framework for assessing the various benefits of data in companies in section 2.3 was validated and strengthened with the empiric data. Several interviewees explicitly divided the benefits of data into operational efficiency (or equivalent) and opportunities to generate new business. While both operational and strategic benefits can be realized with proper use of analytics, those tools enable only internal benefits. In order to tap into external benefits, data must be shared with external parties.

As expected, the connection between the technologies and benefits of data did not come up unprompted in the interviews. However, the interviewees often referred to technologies implicitly, e.g. internet repositories when describing what drove the improvements in standardization processes and when discussing the proposed framework, no contradictions were made. Hence, the proposed link between data sharing technologies and the possibility to exploit the benefits of data remains a viable hypothesis — further research is recommended, as this study has only aimed to develop a viable theory.

Strategic opportunities

Several interviewees acknowledged the existence of a dimension (referred to as *platform* or *ecosystem* opportunities, or equivalent) of benefits that companies in their experience have yet to grasp — they are talked about and opportunities are explored, but very little concrete has been done. These data economy and platform discussions also highlighted how the opportunities lie beyond the company borders; hence this dimension can be considered as an external sub-dimension of strategic opportunities. Some companies might be hesitant to create new products and services because their impact on the company's overall revenues might be insignificant while the development would tie up important resources — giving external companies and individual developers access to the same resources could result in product and service innovation developed outside the company itself without tying up resources, and eventually the aggregated revenues could be significant even for a larger company.

As for other strategic opportunities, they largely revolve around being able to meet customer requirements more thoroughly enabling more revenue opportunities; companies are already doing this to an extent by either purchasing analytics and data, or having similar in-house functions, that allow them to e.g. create more thorough customer profiles or successfully implement dynamic pricing models. However, these benefits are realized primarily by being able to uncover existing internal potential, and can hence be considered as the internal sub-dimension of strategic opportunities. This is also the area in which companies have experienced the most significant growth in recent years and companies are increasingly including relevant external open APIs in their internal data models.

Operational efficiencies

The interviews suggest that operational efficiencies are, indeed, the most important driver for all data use and sharing. The exact drivers vary from "following the industry trends" to "looking to streamline operations". This is well in line with the literature which has largely focused on the cost reduction potential in both information technology and information sharing. Recently, the analytics in companies are used for very similar purposes as two decades ago — inventories are optimized and the amount of manual labor in operations is being reduced. With the emergence of IoT and real-time data, operational efficiencies are also realized with e.g. predictive maintenance or more efficient use of assets and resources.

The empirical evidence for the division between internal and external operational efficiencies remains inconclusive. Although two of the interviewees agreed with the framework and were able to provide examples for internal and external operational efficiencies, they never came up unprompted in the interviews. This could be due to the interviewees themselves as there were no experts from production or supply chain management, which are the typical areas for external operational efficiencies according to the literature. However, there was no contradictory evidence either — the validation of internal and external sub-dimensions of operational efficiencies warrants further research from another study.

Managerial implications

Managers across industries have likely already recognized the various benefits enabled by having access to sufficient data; costs can be cut by streamlining and optimizing operations and revenues can be grown by being able to predict customer needs and behavior. Hence, the internal benefits are already acknowledged and, provided the company has the necessary expertise, they can be brought even further. This study provides examples of benefits companies should be able to expect from sophisticated practices in utilizing the available (or sourced) data and analytics.

The external benefits, both strategic and operational, can only be achieved if the company is willing to share the data externally as well. Operational efficiencies among partners and value chain are already exploited to some extent, as data is shared selectively (e.g. inventory data in manufacturing), and hence cost reductions have been available for some time already. Being able to capitalize on the external strategic opportunities requires companies to share data much more widely and openly — it is risky, however, as the benefits can be unpredictable and such initiatives require sufficient understanding regarding platforms and ecosystems to create comprehensive data strategies in companies. APIs should be considered as they are becoming increasingly relevant across industries, by either regulation or by choice, and proper API strategy is crucial to tap into external strategic opportunities.

Bibliography

Ahmad, S. and Schroeder, R. G. The Impact of Electronic Data Interchange on Delivery Performance. *Production and Operations Management*, 10(1): 16–31, 2001. ISSN 15577813. doi: 10.1080/08853134.1994.10753995.

Alavi, M. and Leidner, D. Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS quarterly*, pages 107–136, 2001. URL `http://www.jstor.org/stable/3250961`.

Antonio, K. W. L., Yam, R. C. M., and Tang, E. The impacts of product modularity on competitive capabilities and performance: An empirical study. *International Journal of Production Economics*, 105:1–20, 2007. doi: 10.1016/j.ijpe.2006.02.002.

Aronson, J. A pragmatic view of thematic analysis. *The qualitative report*, 2(1):1–3, 1995. URL `http://nsuworks.nova.edu/tqr/vol2/iss1/3/?utm{_}source=nsuworks.nova.edu{%}2Ftqr{%}2Fvol2{%}2Fiss1{%}2F3{&}utm{_}medium=PDF{&}utm{_}campaign=PDFCoverPages`.

Arunachalam, V. EDI: An analysis of adoption, uses, benefits and barriers. *Journal of business research*, 46(2):60, 1995.

Arunachalam, V. Electronic Data Interchange: Issues in Adoption and Management. *Information Resources Management Journal*, 10(2):22–31, 1997.

Auramo, J., Kauremaa, J., and Tanskanen, K. Benefits of IT in supply chain management: an explorative study of progressive companies. *International Journal of Physical Distribution & Logistics Management*, 35(2):82–100, 2005.

Baihaqi, I. and Sohal, A. S. The impact of information sharing in supply chains on organisational performance: An empirical study. *Production Planning and Control*, 24(8-9):743–758, 2013. doi: 10.1080/09537287.2012.

666865. URL `http://www.scopus.com/inward/record.url?eid=2-s2.0-84883498567{&}partnerID=40{&}md5=53a91c69ac93b9f4af41e082fffdef23`.

Banerjee, S. and Golhar, D. Y. Electronic data interchange: Characteristics of users and nonusers. *Information and Management*, 26(2):65–74, 1994. ISSN 03787206. doi: 10.1016/0378-7206(94)90054-X.

Barratt, M. and Oke, A. Antecedents of supply chain visibility in retail supply chains: A resource-based theory perspective. *Journal of Operations Management*, 25(6):1217–1233, 2007. ISSN 02726963. doi: 10.1016/j.jom.2007.01.003.

Bednarz, A. Internet EDI: Blending old and new. *Network World*, 21(8): 29–30, 2004.

Benzell, S., LaGarda, G., and Van Alstyne, M. W. The Impact of APIs in Firm Performance. *Ssrn*, pages 1–38, 2017. doi: 10.2139/ssrn.2843326.

Boudreau, K. Open Platform Strategies and Innovation: Granting Access vs. Devolving Control. *Management Science*, 56(10):1849–1872, 2010. ISSN 0025-1909. doi: 10.1287/mnsc.1100.1215.

Brynjolfsson, E. and Hitt, L. M. Beyond Computation: Information Technology. *Journal of Economic Perspectives*, 14(4):23–48, 2000.

Brynjolfsson, E. and McAfee, A. *The Second Machine Age*. W. W. Norton & Company, Inc, New York, NY, 2014.

Brynjolfsson, E., Hitt, L. M. L. M., and Yang, S. Intangible Assets: Computers and Organizational Capital. *Brookings Papers on Economic Activity*, 2002(1):137–198, 2002. ISSN 0007-2303. doi: 10.1353/eca.2002.0003.

Brynjolfsson, E., Hitt, L. M., and Kim, H. H. Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance? *Ssrn*, 2011. doi: 10.2139/ssrn.1819486.

Bunge, J. Big data comes to the farm, sowing mistrust. *Wall Street Journail*, 2014.

Bury, S. Piggly Wiggly's doing it. *Manufacturing Business Technology*, 23 (2):42–44, 2005.

Cachon, G. P. and Fisher, M. Supply Chain Inventory Management and the Value of Shared Information. *Management Science*, 46(8):1032–1048, 2000.

Campbell, D. and Fiske, D. Convergent and Discriminant Validation. *Psychological Bulletin*, 56(2):81–105, 1959.

Chen, H., Storey, V., and Chaing, R. Business Inteligence and Analytics: From Big Data to Big Impact. *MIS Quarterly*, 36(4):1165–1188, 2012.

Chung, K. H., McInish, T. H., Wood, R. A., and Wyhowski, D. J. Production of information, information asymmetry, and the bid-ask spread: Empirical evidence from analysts' forecasts. *Journal of Banking and Finance*, 19(6): 1025–1046, 1995. ISSN 03784266. doi: 10.1016/0378-4266(94)00068-E.

Constantiou, I. D. and Kallinikos, J. New games, new rules: Big data and the changing context of strategy. *Journal of Information Technology*, 30 (1):44–57, 2015. ISSN 14664437. doi: 10.1057/jit.2014.17. URL `http://dx.doi.org/10.1057/jit.2014.17`.

Crum, M. R., Premkumar, G., and Ramamurthy, K. An Assessment of Motor Carrier Adoption, Use, and Satisfaction with EDI. *Transportation Journal (American Society of Transportation & Logistics Inc)*, 35(4):44–75, 1996.

Davenport, T. H. Competing on Analytics. *Harvard Business Review*, (January), 2006.

Davenport, T. H. and Kudyba, S. Designing and Developing Data Products Designing and Developing Data Products. *MIT Sloan Management Review*, 58(1):1–20, 2016.

Dawes, S. S. Interagency Information Shoring: Expected Benefits, Manageable Risks. *Journal of Policy Analysis and Management*, 15(3): 377–394, 1996. URL `http://content.ebscohost.com/ContentServer.asp?T=P{&}P=AN{&}K=9608131199{&}S=R{&}D=bth{&}EbscoContent=dGJyMNLr40Seqa44zdnyOLCmr1Cep65SsKq4SraWxWXS{&}ContentCustomer=dGJyMPGptkyvq7ZJuePfgeyx44Dt6fIA`.

Dubois, A. and Gadde, L. E. Systematic combining: An abductive approach to case research. *Journal of Business Research*, 55(7):553–560, 2002. ISSN 01482963. doi: 10.1016/S0148-2963(00)00195-8.

Emmelhainz, M. A. *EDI: A Total Management Guide*. John Wiley & Sons, Inc., New York, 1992.

Fitzgerald, M. An Internet for Manufacturing. *MIT Technology Review*, 2013. URL `https://www.technologyreview.com/s/509331/an-internet-for-manufacturing/`.

Fosso Wamba, S., Akter, S., Edwards, A., Chopin, G., and Gnanzou, D. How 'big data' can make big impact: Findings from a systematic review and a longitudinal case study. *International Journal of Production Economics*, 165:234–246, 2015. ISSN 09255273. doi: 10.1016/j.ijpe.2014.12.031. URL `http://dx.doi.org/10.1016/j.ijpe.2014.12.031`.

Gavirneni, S., Kapuscinski, R., and Tayur, S. Value of information in capacitated supply chains. *Management Science Publication*, 45(1):16–24, 1999.

Gawer, A. *Platforms, markets and innovation.* Edward Elgar Publishing, 2009.

Gawer, A. Bridging different perspectives on technological platforms: Toward an intergrative framework. *Research Policy*, 43:1239–1249, 2014.

Gawer, A. and Cusumano, M. Platform Leaders. *MIT Sloan Management Review*, Special Co(Top 10 Lessons on Strategy):68–75, 2008.

Ghazawneh, A. *Towards a Boundary Resources Theory of Software Platforms.* Number 85. 2012. ISBN 9789186345389. doi: 10.1016/S0009-2614(95)01235-4. URL `http://www.diva-portal.org/smash/record.jsf?pid=diva2{%}3A567769{&}dswid=1262`.

Ghazawneh, A. and Henfridsson, O. Governing Third-Party Development through Platform Boundary Resources. *ICIS 2010 Proceedings*, pages 1–18, 2010. URL `http://aisel.aisnet.org/icis2010{_}submissions/48`.

Greene, J. C., Caracelli, V. J., and Graham, W. F. Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis*, 11(3):255, 1989. ISSN 01623737. doi: 10.2307/1163620. URL `http://links.jstor.org/sici?sici=0162-3737{%}28198923{%}2911{%}3A3{%}3C255{%}3ATACFFM{%}3E2.0.CO{%}3B2-F{&}origin=crossref`.

Günther, W. A., Rezazade Mehrizi, M. H., Huysman, M., and Feldberg, F. Debating big data: A literature review on realizing value from big data. *Journal of Strategic Information Systems*, 26(3):191–209, 2017. ISSN 09638687. doi: 10.1016/j.jsis.2017.07.003.

Henderson, R. M. and Clark, K. B. Architectural Innovation: The Reconfiguration of Existing Product Technologies and the Failure of Established Firms. *Administrative Science Quarterly*, 35:9–30, 1990.

Huang, Z., Janz, B. D., and Frolick, M. N. A comprehensive examination of Internet-EDI adoption. *Information Systems Management*, 25(3):273–286, 2008. ISSN 10580530. doi: 10.1080/10580530802151228.

Iacovou, C. L., Benbasat, I., and Dexter, A. S. Electronic Data Interchange and Small Organizations: Adoption and Impact of Technology. *MIS Quarterly*, 19(4):465, 2006. ISSN 02767783. doi: 10.2307/249629.

Iskandar, B. Y., Kurokawa, S., and LeBlanc, L. J. Adoption of electronic data interchange: The role of buyer-supplier relationships. *IEEE Transactions on Engineering Management*, 48(4):505–517, 2001. ISSN 00189391. doi: 10.1109/17.969427.

Iyer, B. and Subramaniam, M. The Strategic Value of APIs. *Harvard Business Review*, 2015. doi: .1037//0033-2909.I26.1.78.

Jacobson, D., Brail, G., and Woods, D. *APIs: A strategy guide*. O'Reilly Media, Inc, 2011.

Jick, T. D. Mixing Qualitative and Quantitative Methods: Triangulation in Action. *Administrative Science Quarterly*, 24(4):602, 1979. ISSN 00018392. doi: 10.2307/2392366. URL `https://www.jstor.org/stable/2392366?origin=crossref`.

Jonsen, K. and Jehn, K. A. Using triangulation to validate themes in qualitative studies. *Qualitative Research in Organizations and Management: An International Journal*, 4(2):123–150, 2009. ISSN 17465656. doi: 10.1108/17465640910978391.

Joseph, S., Ludford, V., and McAllister, B. Plugging in: Enabling the enterprise for the platform economy. *Technical report, Gartner Research Board*, 2016.

Katz, M. L. and Shapiro, C. Systems Competition and Network Effects. *Journal of Economic Perspectives*, 8(2):93–115, 1994. ISSN 0895-3309. doi: 10.1257/jep.8.2.93.

Lau, J. S., Mak, K., and Huang, G. The impacts of sharing production information on supply chain dynamics: A review of the literature. *International Journal of Production Research*, 41(7):1483–1517, 2003. URL `http://www.scopus.com/inward/record.url?eid=2-s2.0-0242364629{&}partnerID=40`.

Lee, H. L. The Triple-A Supply Chain. *Harvard Business Review*, 82(10):102–112, 2004.

Lee, H. L. and Whang, S. Information Sharing in a Supply Chain. *International Journal of Technology Management*, 20(3-4):373–387, 2000.

Lee, H. L., Padmanabhan, V., and Whang, S. Information Distortion in a Supply Chain: The Bullwhip Effect. *Management Science*, 43(4):546–558, 1997. ISSN 0025-1909. doi: 10.1287/mnsc.43.4.546. URL `http://pubsonline.informs.org/doi/abs/10.1287/mnsc.43.4.546`.

Lee, H. L., So, K. C., and Tang, C. S. The Value of Information Sharing in a Two-Level Supply Chain. *Management Science*, 46(5):626–643, 2000. ISSN 0025-1909. doi: 10.1287/mnsc.46.5.626.12047.

Lee, S. and Lim, G. G. The impact of partnership attributes on EDI implementation success. *Information and Management*, 41:135–148, 2003. ISSN 03787206. doi: 10.1016/S0378-7206(03)00153-8.

Lewis, M. W. Iterative triangulation: a theory development process using existing case studies. *Journal of Operations Management*, 16:455–469, 1998.

Liozu, S. M. and Ulaga, W. *Monetizing Data*. Value Innoruption Advisors Publishing, Arizona, 2018. ISBN 978-1-945815-04-1.

Lotfi, Z., Mukhtar, M., Sahran, S., and Zadeh, A. T. Information sharing in supply chain management. *Procedia Technology*, 11:298–304, 2013. ISSN 2212-0173. doi: 10.1109/DCABES.2010.89. URL `http://dx.doi.org/10.1016/j.protcy.2013.12.194`.

Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., and Hung Byers, A. Big data: The next frontier for innovation, competition and productivity. *McKinsey Global Institute*, 2011.

Mathison, S. Why Triangulate? *Educational Researcher*, 17(2):13–17, 1988.

Mayer-Schönberger, V. and Cukier, K. *Big Data - A Revolution That Will Transform How We Live, Work and Think*. John Murray, 2013.

McAfee, A. and Brynjolfsson, E. Big Data: The Management Revolution. *Harvard Business Review*, (October), 2012. doi: 10.3823/5004.

Miles, M. and Huberman, A. *Qualitative data analysis*. Sage, 1984.

Modell, S. Triangulation between case study and survey methods in management accounting research: An assessment of validity implications. *Management Accounting Research*, 16(2):231–254, 2005. ISSN 10445005. doi: 10.1016/j.mar.2005.03.001.

Moilanen, J., Niinioja, M., Seppänen, M., and Honkanen, M. *API Economy 101.* Books on Demand, Helsinki, 2019.

Mukhopadhyay, T., Kekre, S., and Kalathur, S. Business Value of Information Technology: A Study of Electronic Data Interchange. *MIS Quarterly*, pages 137–156, 1995.

Najjar, M. S. and Kettinger, W. J. Data Monetization: Lessons from a retailer's journey. *MIS Q. Exec.*, 12(4):12, 2013. ISSN 15401960.

Narayanan, S., Marucheck, A. S., and Handfield, R. B. Electronic data interchange: Research review and future directions. *Decision Sciences*, 40 (1):121–163, 2009. ISSN 00117315. doi: 10.1111/j.1540-5915.2008.00218.x.

Nunes, J. C. and Drèze, X. Your Loyalty Program Is Betraying You. *Harvard Business Review*, 84(4), 2006. URL `www.hbr.org`.

Parker, G. and Van Alstyne, M. W. Innovation, Openness, and Platform Control. *Management Science*, 64(7):3015–3032, 2018. ISSN 0025-1909. doi: 10.2139/ssrn.1079712.

Parker, G., Van Alstyne, M. W., and Jiang, X. Platform Ecosystems: How Developers Invert the Firm. *Ssrn*, pages 0–30, 2016a. ISSN 02767783. doi: 10.2139/ssrn.2861574.

Parker, G. G., Van Alstyne, M. W., and Choudary, S. P. *Platform revolution: How networked markets are transforming the economyand how to make them work for you.* WW Norton & Company, 2016b.

Peppard, J., Edwards, C., and Lambert, R. Crowdsourcing: How to benefit from (too) many great ideas. *MIS Quarterly Executive*, 10(2):115–117, 2011. ISSN 15401960. doi: 10.1108/02635570910926564.

Pfeiffer, H. K. C. *The Diffusion of Electronic Data Interchange.* Springer-Verlag, New York, NY, 1992.

Porter, M. *Competitive Strategy.* Free Press, New York, NY, 1980.

Porter, M. *Competitive Advantage.* Free Press, New York, NY, 1985.

Prajogo, D. and Olhager, J. Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration. *International Journal of Production Economics*, 135 (1):514–522, 2012. ISSN 09255273. doi: 10.1016/j.ijpe.2011.09.001. URL `http://dx.doi.org/10.1016/j.ijpe.2011.09.001`.

Prosser, A. and Nickl, A. The impact of EDI on interorganizational integration. *International Journal of Production Economics*, 52(3):269–281, 1997. ISSN 09255273. doi: 10.1016/S0925-5273(97)00088-1.

Purvis, R. L., Sambamurthy, V., and Zmud, R. W. The Assimilation of Knowledge Platforms in Organizations. *Organization Science*, 12(2):117–135, 2001. ISSN 1047-7039. doi: 10.1287/orsc.12.2.117.10115.

Rajala, R., Hakanen, E., Mattila, J., Seppälä, T., and Westerlund, M. How Do Intelligent Goods Shape Closed-Loop Systems? *California Management Review*, 60(3):20–44, 2018. ISSN 21628564. doi: 10.1177/0008125618759685.

Reekers, N. Electronic data interchange use in German and US organizations. *International Journal of Information Management*, 14(5):344–356, 1994. ISSN 02684012. doi: 10.1016/0268-4012(94)90072-8.

Robinson, D. G. and Stanton, S. A. Exploit EDI before EDI exploits you. *Information Strategy, The Executive's Journal*, (Spring):32–35, 1987.

Rochet, J.-C. and Jean, T. Platform Competition Two-Sided Markets. *The European Econimics Association*, pages 991–1029, 2003. ISSN 00049018. doi: 10.1111/j.1467-8462.2013.12020.x.

Sanchez, R. and Mahoney, J. T. Modularity, Flexibility, and Knowledge Management in Product and Organization Design. *Strategic Management Journal*, 17(Special Issue):63–76, 1996.

Santos, W. ProgrammableWeb API Directory Eclipses 17,000 as API Economy Continues Surge, 2017. URL `https://www.programmableweb.com/news/programmableweb-api-directory-eclipses-17000-api-economy-continues-surge/research/2017/03/13`.

Schilling, M. A. Toward a General Modular Systems Theory and Its Application to Interfirm Product Modularity. *The Academy of Management Review*, 25(2):312–334, 2000.

Schulz, P., Matthe, M., Klessig, H., Simsek, M., Fettweis, G., Ansari, J., Ashraf, S. A., Almeroth, B., Voigt, J., Riedel, I., Puschmann, A., Mitschele-Thiel, A., Muller, M., Elste, T., and Windisch, M. Latency Critical IoT Applications in 5G: Perspective on the Design of Radio Interface and Network Architecture. *IEEE Communications Magazine*, 55(2):70–78, 2017. ISSN 01636804. doi: 10.1109/MCOM.2017.1600435CM.

Smith, G., Ofe, H. A., and Sandberg, J. Digital service innovation from open data: exploring the value proposition of an open data marketplace. *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2016-March:1277–1286, 2016. ISSN 15301605. doi: 10.1109/HICSS.2016.162.

Star, S. L. and Griesemer, J. R. Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3):387–420, 1989.

Starr, M. Modular production a new concept. *Harvard Business Review*, (November-December):131–142, 1965.

Stefansson, G. Business-to-business data sharing: A source for integration of supply chains. *International Journal of Production Economics*, 75(1-2): 135–146, 2002. ISSN 09255273. doi: 10.1016/S0925-5273(01)00187-6.

Stevens, G. C. Integrating the Supply Chain. *International Journal of Physical Distribution & Materials Management*, 19(8):3–8, 1989. ISSN 0020-7527. doi: 10.1108/EUM0000000000329.

Takac, P. F. Electronic data interchange (EDI): an avenue to better performance and the improvement of trading relationships. *International Journal of Computer Applications in Technology*, 5(1):22–36, 1992.

Tan, K. C., Kannan, V. R., and Handfield, R. B. Supply chain management: supplier performance and firm performance. *International Journal of Purchasing & Materials Management*, 34(3):2–9, 1998.

Tao, F., Qi, Q., Liu, A., and Kusiak, A. Data-driven smart manufacturing. *Journal of Manufacturing Systems*, 48:157–169, 2018. ISSN 02786125. doi: 10.1016/j.jmsy.2018.01.006. URL `https://doi.org/10.1016/j.jmsy.2018.01.006`.

Tapscott, D. and Williams, A. D. *Wikinomics: how mass collaboration changes everything*. Penguin, 2006.

Teece, D. J., Pisano, G., and Shuen, A. M. Y. Dynamic Capabilities and Strategic Management. *Strategic Management Journal*, 18(7):509–533, 1997.

The Economist. Beyond the PC. *Special Report on Personal Technology*, 2011. URL `http://www.economist.com/node/21531109`.

Tiwana, A. Does Technological Modularity Substitute For Control? A Study of Alliance Preformance in Software Outsourcing. 780(April 2007):769–780, 2008. doi: 10.1002/smj.

Tiwana, A. and Konsynski, B. Complementarities Between Organizational IT Architecture and Governance Structure. *Information Systems Research*, 21(2):288–304, 2010.

Tiwana, A., Konsynski, B., and Bush, A. A. Platform evolution: Coevolution of platform architecture, governance, and environmental dynamics. *Information Systems Research*, 21(4):675–687, 2010. ISSN 15265536. doi: 10.1287/isre.1100.0323.

Tuunanen, T. and Cassab, H. Service Process Modularization : Reuse Versus Variation in Service Extensions. *Journal of Service Research*, 14(3):340–354, 2011. doi: 10.1177/1094670511413912.

Vickery, S. K., Jayaram, J., Droge, C., and Calantone, R. The effects of an integrative supply chain strategy on customer service and financial performance: An analysis of direct versus indirect relationships. *Journal of Operations Management*, 21(5):523–539, 2003. ISSN 02726963. doi: 10.1016/j.jom.2003.02.002.

Vijayasarathy, L. R. and Tyler, M. L. Adoption factors and electronic data interchange use: a survey of retail companies. *International Journal of Retail Distribution & Management*, 25:286–292, 1997.

Voss, C. A. and Hsuan, J. Service Architecture and Modularity. *Decision Sciences*, 40(3):541–569, 2009.

Watson, I. Internet, intranet, extranet: Managing the information bazaar. *Aslib Proceedings*, 51(4):109–114, 1999. ISSN 0001253X. doi: 10.1108/EUM0000000006969.

Werner, T. EDI Meets The Internet. *Transportation & Distribution*, 40(6):36–44, 1999. ISSN 1092-4388. doi: 10.1086/250095.

Westbrook, R. and Frohlich, M. Arcs of Integration: An International Study of Supply Chain Strategies. *Journal of Operations Management*, 19:185–200, 2001. URL `http://eureka.bodleian.ox.ac.uk/1829/`.

Wixom, B. H. and Boss, J. W. How to Monetize Your Data. *MIT Sloan Management Review*, 58(3), 2017. doi: 10.7551/mitpress/11633.003.0009.

Woerner, S. L. and Wixom, B. H. Big data: extending the business strategy toolbox. *Journal of Information Technology*, 30:60–62, 2015.

Wu, G., Talwar, S., Johnsson, K., Himayat, N., and Johnson, K. D. M2M: From mobile to embedded internet. *IEEE Communications Magazine*, 49 (4):36–43, 2011. ISSN 01636804. doi: 10.1109/MCOM.2011.5741144.

Yoo, Y., Henfridsson, O., and Lyytinen, K. The new organizing logic of digital innovation: An agenda for information systems research. *Information Systems Research*, 21(4):724–735, 2010. ISSN 15265536. doi: 10.1287/isre.1100.0322.

Yu, Z., Yan, H., and Cheng, T. E. Benefits of information sharing with supply chain partnerships. *Industrial Management and Data Systems*, 101(3):114–119, 2001. ISSN 02635577 (ISSN). URL `http://www.scopus.com/inward/record.url?eid=2-s2.0-0035781053{&}partnerID=40{&}md5=e529ef1f6fed07e0a0088ba6efa5d099`.

Zhu, F. and Liu, Q. Competing with complementors: An empirical look at Amazon.com. *Strategic Management Journal*, 39(10):2618–2642, 2018. ISSN 10970266. doi: 10.1002/smj.2932.

Zhu, K., Kraemer, K. L., Gurbaxani, V., and Xu, S. Migration to Open-Standard Interorganizational Systems: Network Effects, Switching Costs, and Path Dependency. *MIS Quarterly*, 30:515–539, 2006. ISSN 02767783. doi: 10.2307/25148771.

Appendix A

Interview guide

Evolution of data sharing mechanisms

- What is the earliest form of information exchange / data sharing outside of company borders that you know of?
  - How did it work / where was it used?
  - What kind of data was shared at that time? Why?
  - What was the reason for adopting this practice and what were expected benefits? Were they realized according to the expectations?
    - Were there any unforeseen benefits of either the technology or the data sharing practices?
  - Is this technology still used in the company? Why?
- After (1), how did data sharing practices and technologies evolve?
  - What were the key adoptions the company made in regard to either interorganizational information systems or data sharing practices?
    - How did they work / where were they used?
    - What were the reasons for adopting these practices and what were the expected benefits? Were they realized according to the expectations?
      - Were there any unforeseen benefits of either the technology or the data sharing practices?
    - Are these technologies or practices still used in the company? Why?

- How did the partners with whom data was shared change with these technologies or practices?
  - Did these new partnerships change create new business opportunities, new products or new services?
    - If yes, who was able to capitalize on it?
    - How long did it take to realize?
    - Are there any examples?
- Considering the last three decades, what would you perceive as the overarching trends in data sharing practices with regards to technologies, variety of data shared, the number of partners, and the sought / realized benefits?
- Specifically, if applicable, how did the emergence of Internet change data sharing practices?
- Specifically, if applicable, how has the emergence of mobile computing and increased connectivity changed data sharing practices?

Data sharing and trading

- Does your company trade data? If yes;
  - What is its impact on revenue / profit?
  - For how long has it been exercised?
- Generally, how would you describe the benefits of data [sharing]?
  - What kind of dimensions would there be?
  - What would be the variables to measure the monetary value of said benefits?
  - Are they same to all companies?
- What are the other options of doing business with data? How is it used, how could it be used more, how could it be monetized?
- How vital is data sharing in your business? Why?
  - How much of revenue is dependent or linked to data sharing?
  - Do you think your answer could be generalized across your company's industry? Why?

Theoretical framework

- Given your experience in business, does this framework reflect your experiences?
  - Why / why not?
- Are you able to place the aforementioned steps of the evolution within this framework?
  - Given the theoretical benefits of data sharing at these stages, which ones were realized? Which ones were not? What is missing?
- How far in this evolution is your company?
  - What are the causes of friction?
  - What are the causes of adoption?
- Given the framework, what kind of opportunities could you hypothetically see in your company if the data sharing practices and technologies were developed further?